Defining Approaches and Tools for the Creation of Shareable Datasets

Confirmation Report

June 2010

Heather McDonald (swan-h2@email.ulster.ac.uk)
Computer Science Research Institute
School of Computing and Mathematics
Faculty of Computing and Engineering
Room 16J26
University of Ulster, Jordanstown

Supervisors:
- Prof. Chris Nugent
- Dr. George Moore
- Dr. Dewar Finlay
Table of Contents

1.0 Introduction 4
   1.1 Ageing Population 4
   1.2 Chronic Diseases 5

2.0 Assistive Technologies 6
   2.1 Smart Home Environments 7
   2.2 Assistive Devices 8
   2.3 Leading Smart Home Research Groups 10
      2.3.1 Gator Tech Smart House, University of Florida 10
      2.3.2 Aware Home, Georgia Institute of Technology 11
      2.3.3 PlaceLab, Massachusetts Institute of Technology 11
      2.3.4 Duke Smart Home, Duke University 11
      2.3.5 Tiger Place, University of Missouri 11
      2.3.6 MavHome, University of Texas, Arlington 12
   2.4 Problem Identified 12

3.0 Data Management 12
   3.1 Data Formatting 12
      3.1.1 XML 13
      3.1.2 XML Schema 13
   3.2 Data Storage 13
      3.2.1 XML Enabled Database v Native XML Database 14
   3.3 Ontologies 15
   3.4 Similar Data Storage and Formatting Projects 15
      3.4.1 Approaches in Healthcare 15
      3.4.2 Approaches related to Smart Homes 17
      3.4.3 Key Research Projects 19

4.0 Contribution 20

5.0 Work to Date 21

6.0 Future Work 27

7.0 Conclusion 29
8.0 References

Appendix A – homeML version 1.0 review
Appendix B – homeML version 1.0 hierarchical tree diagram
Appendix C – Scenarios
Appendix D – homeML version 2.0 review
Appendix E – Proposed Work Schedule
Appendix F – Thesis Outline
Appendix G – Completed Research Training Seminars
Appendix H – ITAB’10 Paper
Appendix I – NIBES’10 Abstract
Abstract

homeML is an XML-based format designed to store data generated within smart home environments. It was originally proposed as a means of solving the issues caused by the heterogeneous nature of data recorded within such environments. In our current work we have extended upon the concepts of homeML and have subsequently developed an improved format to support the storage and exchange of sensory data collected from heterogeneous sources both inside and outside of the home environment. Within this report an XML-based format has been proposed and successfully tested by storing datasets generated by five research institutions.

1.0 Introduction

1.1 Ageing Population

As the population continues to grow the percentage of elderly is also increasing. Studies have shown that many populations are experiencing a shift from a largely young population to a population with a greater proportion of elderly people [1]. The proportion of the global population with an age greater than 60 is expected to increase from 10.0% in 2000 to 21.8% in 2050 and reach 32.2% by 2100 [2]. In 1948 the percentage of the UK population aged between 65 and 74 years was 7.2%, this has since increased to 8.2% in 2006 and is expected to reach 10.7% by 2051 [3], as shown in Figure 1. In addition to this the average age of life expectancy has grown from 75.7 years in 1990 to 79.7 years in 2008 and is expected to further increase [4].

![Figure 1. Estimated Percentage of People over 65 years living within the United Kingdom [3]](image)

The number of children being born is also predicted to decrease, with the ratio of 15 to 65 year olds expected to decrease from 9:1 to 4:1 by 2050 [5]. As a result, there will be fewer young people to assist older adults cope with the challenges of ageing in the future [6].
As people grow older they are more likely to experience age-related impairments; with the older population being the most common sufferers of disease, co-morbidity and disabilities [1], resulting in a major impact on the economy [7].

Although remaining at home can become challenging for an elderly person; over 80% of people over 65 years old manage to continue to live at home as they age [1]. Research has also shown that patients discharged from hospital to a nursing home live on average 2 1/2 years less than someone the same age, discharged from hospital at the same time but returning home [8].

As people age they may begin to experience challenges at home, however, it is not unreasonable for these people to want to remain at home and maintain their independence. Therefore, ageing-in-place refers to members of the elderly community continuing to live where they have lived for years, typically not in a healthcare environment, using products and services which allow them to remain at home as their circumstances change [9].

Due to these significant demographic changes there is an increasing need to re-assess the current approaches to health and social care provision, particularly in the direction of health, technology and the home [1]. Therefore, enabling the enhancement of the elderly populations overall quality of life whilst allowing them to maintain their independence.

1.2 Chronic Diseases

Coupled with population growth will be the prevalence of the increasing numbers of people suffering from a chronic disease. Chronic diseases typically affect a patient over a long period of time and are one of the major causes of death and disability globally, representing 60% of all deaths [10]. There are a large number of chronic diseases including dementia, cancer, stroke, diabetes and cardiovascular disease. In the United Kingdom alone, chronic diseases account for 85% of all deaths [11]; currently the cost to the government is £12 billion a year, with this bill expected to rise to £15.6 billion by 2040[12].

As a result of these figures the UK government have introduced self-management approaches, providing patients with the ability to manage their condition, treatment and symptoms [13]. Within the last 10 years the UK government have developed an Expert Patient Programme (EPP); the NHS approach to chronic disease management [14]. An expert patient is someone who feels positive and in control of their life, who is realistic about the affect their condition has on them and their family and who feels confident discussing their condition with healthcare professionals [15]. The programme was introduced in 2002 and involves a training course teaching people how to manage their condition using 5 core skills [16]:

1. Problem solving
2. Making the best use of resources
3. Decision making

4. Taking appropriate action

5. Building better relationships between the patient and their care team

Technology plays a major part in enabling patients to manage their condition. There are a wide range of technology solutions available to collect data, educate the patient on their condition and communicate with the patient care team [17]. The technologies can be categorised as [17]:

- **Patient Technologies** – Devices located on the patient in order to collect data.
- **Everyday Technologies** – Non-medical devices such as mobile phones.
- **Medical Devices** – Devices used to collect physiological data that can transferred to a provider as part of a care management program.
- **Digital Medical Home** – Devices and systems embedded into the home environment.

### 2.0 Assistive Technologies

Assistive Technologies are “any item, piece of equipment, or product system, whether acquired commercially off the shelf, modified or customised, that is used to increase, maintain or improve the functional capabilities of individuals with disabilities” [18]. Smart applications are now being employed to monitor, manage, and motivate health behaviours both inside and outside of the home on a constant basis, as shown in Table 1 [19].

<table>
<thead>
<tr>
<th>Function</th>
<th>Selected Technologies</th>
<th>Innovation and Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Monitor</strong></td>
<td>Continuous monitoring of a patient's health and the activities they perform</td>
<td>Rapid intervention should the patient experience a change in their condition, quick response to emergencies</td>
</tr>
<tr>
<td><strong>Manage</strong></td>
<td>Identify patients who require a healthcare intervention</td>
<td>Improved delivery of healthcare</td>
</tr>
<tr>
<td><strong>Motivate</strong></td>
<td>Educate a patient on their condition and motivate them.</td>
<td>Management of current condition as well as prevention of additional diseases.</td>
</tr>
</tbody>
</table>

SOURCE: MIT AgeLab.

**Table 1.** A summary of technology use in the home, adapted from [19].
2.1 Smart Home Environments

The term ‘smart home’ refers to residences equipped with technologies that allow the monitoring of its inhabitants; whilst encouraging independence and the maintenance of health and well being [5]. Due to the demographic changes previously discussed, significant effort has been directed towards research in the area of home based support and health care provision. This has resulted in an array of solutions becoming available which can be used to monitor and support people within and outside of their homes. In addition many types of assistive technologies have been developed to augment multiple activities of a person’s day-to-day life.

Ambient Assisted Living (AAL) can help the elderly population live longer at home, whilst ensuring a high quality of life, through the application of ICT combined with intelligent software and devices [20]. The European Commission are currently funding the AAL Joint Programme, which aims to improve people's quality of life whilst concurrently improving the industrial base of ICT within Europe [21]. In order to achieve this aim the AAL Joint Programme provides significant funding to assist research and development projects in the AAL domain [21].

As well as smart home environments being used by people suffering from cognitive or physical impairments, they can also support those people who may live alone and may require assistance in an emergency. In addition, people who receive healthcare at a distance or people who suffer long term chronic diseases and require continuous monitoring may also benefit.

Cooke et al. [22] have identified the typical components of a smart home environment, as shown in Figure 2. Sensors monitor an environments physical components, which enables agents to reason about the environment and select an action using the data collected, which will then be performed by actuators [22].
2.2 Assistive Devices

As each person can have varying needs, it is essential that any support being delivered suits their individual needs and capabilities[5]. Devices can be either wearable, implantable or microcapsule and can be either worn by the individual or implanted into the home [5]. The devices will then access various parameters including body motion, vital signs, social interactions and ambience (including light and room temperature) in addition to many more [5].

Table 2 [23] provides a sample of devices, along with their potential uses, that may be employed by individuals living within a smart home environment.
<table>
<thead>
<tr>
<th>Device</th>
<th>Purpose</th>
</tr>
</thead>
</table>
| PIR Movement Sensor            | Should be wall mounted within each room of a dwelling to monitor movement. Typically the sensor will have 2 modes:  
**Activity Mode** – An alarm will sound should there be long periods of inactivity.  
**Intruder Mode** – An 'intruder' alarm will sound when movement is detected. |
| Flood Detector                 | Can be placed anywhere there is a risk of flooding such as the bathroom or the kitchen. Should moisture be detected an alarm will sound notifying the resident. |
| Temperature Extremes Sensor    | Should be wall mounted and will sound an alarm should it detect any extreme temperature changes that could result in hypothermia or dehydration. |
| Smoke Detector                 | Should be ceiling mounted and will sound an alarm should it detect smoke.                         |
| Radio Door Contacts            | Should be mounted to a door and corresponding door frame to record when a door has been opened. Radio Door Contacts can be used to monitor an individual’s movement within their environment. |
| Fall Detector                  | Falls will be detected and an alarm sounded. As well as this, stumbles and recovered falls can be recorded and used as a means of monitoring a patient’s condition. |
| Bed/Chair Sensor               | The sensor will detect a person has got up from the bed or a chair and will sound an alarm should the person not return within a preset time. |
| Blood Pressure Monitor         | Allows the individual to monitor their blood pressure levels from home without the need for GP/Carer assistance. |
| Pulse Oximeter Sensor          | Allows the individual to monitor their blood/oxygen levels from home without the need for GP/Carer assistance. |

*Table 2.* Sample of smart devices currently available within the home [23].

*Table 3* provides a sample of actuators, which can be integrated into the smart home environment in order to perform actions in response to data gathered from the sensors documented in *Table 2*.

<table>
<thead>
<tr>
<th>Device</th>
<th>Time Based Control</th>
<th>Optimiser Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heating</td>
<td>Will turn on and off the home heating system at pre-selected times [24].</td>
<td>Will generally involve turning on the heating system when external temperatures reach a pre-defined temperature [24].</td>
</tr>
<tr>
<td>Lighting</td>
<td>Will turn on and off lights within the home at pre-selected times [24].</td>
<td>Will generally involve turning on or off lights to achieve a pre-defined aluminous level [24].</td>
</tr>
<tr>
<td>Door Openers</td>
<td>Doors will automatically open when an individual approaches [25].</td>
<td></td>
</tr>
</tbody>
</table>

*Table 3.* Sample of actuators currently available within the home.
Table 4 provides a sample of two devices that may be employed by individuals for constant monitoring independent of their environment. This allows the individual to be monitored both inside and outside of the home.

<table>
<thead>
<tr>
<th>Smart Garments</th>
<th>Commercially there are a number of garments available that can be worn by an individual to continuously record vital signs [26].</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smart Phones</td>
<td>Smart phones can be used to perform various tasks such as reminders and alarms, as well as this they are also capable of being used as a pedometer (counts number of steps taken) and GPS (Global Positioning System) as a means of tracking the movement of an individual outside of their home environment. Such features are available with the new generation of mobile phones, an example of which being the Apple iPhone [27].</td>
</tr>
</tbody>
</table>

Table 4. Sample of smart device currently available for constant monitoring outside of the home.

Sensors inside and outside of the home environment monitor physical components in order to produce data. This data is typically stored within a database where it can be processed further in order to produce useful knowledge. Processing can involve filtering data, comparing datasets or applying algorithms to the data in order to recover knowledge, allowing behavioural patterns to be identified [22]; which can then be used to diagnose conditions, as well as monitor and identify cognitive or physical decline.

2.3 Leading Smart Home Research Groups

Within the last decade organisations from both the technical and health and social care perspective have sought to develop technologies that can be used to monitor, diagnose and manage patients’ conditions inside and outside of their home environment. This has been motivated by the requirement to move from reactive care provision to proactive care provision in addition to home based care as opposed to hospital care. The section below documents six Smart Home Projects; however, there are many more including Microsoft's EasyLiving Project [28], CarerNet, UK [29] and the Elite CARE Project, Portland, Oregon [30].

2.3.1 Gator-Tech Smart House, University of Florida

The Gator-Tech Smart House, is the result of over than 10 years of research, with the aim of developing a home environment, that can assist its inhabitants [31]. The Gator-Tech house has a number of integrated sensors and devices, such as smart mirrors that will display messages and reminders, a smart floor that uses sensors to track the inhabitants’ movement and a smart phone that not only has the traditional functionality of a telephone but can also be used to control appliances within the environment. As well as this the house also incorporates actuators, some of which are devices that the inhabitant can interact with, such as turning on a lamp or heater.
2.3.2 Aware Home, Georgia Institute of Technology

The Aware Home/Georgia Tech Broadband Institute Residential Laboratory is a 3 storey home that is currently being used as a laboratory to support developments from design to evaluation [32]. Currently the Aware home supports several scenarios ranging from busy families to the elderly with various embedded devices and applications. Applications available include a Family Video Archive [32] that will display videos to the user and a Technology Coach [32] that provides elderly users with feedback when using devices embedded within the home.

2.3.3 PlaceLab, Massachusetts Institute of Technology

PlaceLab is a real home where research volunteers can live for a period of time. The participants perform routine activities and interactions that are observed and recorded for later analysis [33]. Sensors are embedded into the home to record descriptions of activities in order to generate datasets of typical domestic activities. The datasets can then assist research in pervasive computing, preventative healthcare and user interface design [33].

2.3.4 Duke Smart Home, Duke University

The Duke Smart Home/The Home Depot Smart Home is a 6,000 square foot research residence. Its aim is to make a contribution in future residential building technology, concentrating in the area of smart home environments [34]. The project aims to "explore an energy efficient lifestyle; compare, use and develop smart and sustainable technology and provide an insight into homeowners do-it-yourself technology integration and control" [34].

2.3.5 Tiger Place, University of Missouri

Tiger Place is a retirement home, who emphasis is on health, wellness and preventive care (TigerCare); allowing inhabitants to receive healthcare in the comfort of their own home [35]. Tiger Place offers a wide variety of lifestyle options including Independent or Assisted Living and Alzheimer's Care [35].

2.3.6 MavHome, University of Texas, Arlington

The MavHome project is focused upon the creation of intelligent home environments, viewing an intelligent environment as consisting of sensors and actuators that respond to changes in the environment [36]. The aim of the project is to minimise the cost of maintaining the home whilst maximising the residents comfort; through the ability of the house to reason and adapt to its environment [36].

2.4 Problem Identified

Due to the high level of interest in smart homes many organisations from both the technical and health and social care perspectives have sought to create a range of solutions which have subsequently
resulted in, an abundance of data being generated [37]. The data generated as a direct result of this is heterogeneous, given that it is gathered from a number of sources and services and subsequently it is stored in various formats.  

The main issue with the heterogeneity of data generated within a smart home environment is its lack of interoperability when data is being processed by a number of services. There are a range of protocols enabling the integration of numerous devices and services within a smart home environment such as KONNEX (the only global standard for building and home control [38]). KONNEX involves the use of a communication system to connect numerous and varied devices within a home or building [39]. Nevertheless, to date, there does not exist a standard within the research community specifying how data generated within a smart home environment should be stored. Due to the issues caused by the heterogeneous nature of the data there is a growing need to develop a standard format.

3.0 Data Management

Data management can be defined as “the development, execution and supervision of plans, policies, programs and practices that control, protect, deliver and enhance the value of data and information assets” [40].

3.1 Data Formatting

Data formatting refers to the organisation of data according to preset specifications [41]. Consequently, through the creation of a data standard, specifying how data generated within smart home environments should be stored; issues arising as a direct result of the data being of a heterogeneous nature.

3.1 1 XML

eXtensible Markup Language (XML) is a rule based language designed to store and transfer data. It is one of the most common tools for data transmission and is becoming increasingly popular when describing and storing information [42]. It was designed to generate models and describe data by the use of an XML schema or a Document Type Definition (DTD) [43]. Although XML may appear plain and simple, ‘it is a multifaceted concept defying oversimplified definitions’ [44]. As XML is a mark-up language it adds labels to the content in order to enhance the content in some way [44].

The main benefits of using XML are its application, vendor and platform independence; whilst following a straightforward hierarchical data structure [37]. XML is also considered to be simple to read and understandable by both users and computers. In addition, XML is also capable of unifying

---

1 Here, format refers to data being saved according to a preset specification.
varied forms of information, as it can be easily parsed, edited and can be used to represent both semi-structured and structured data [45].

Finally, an increasing number of people wish to access and manipulate useful information from a multitude of data sources and achieve interoperability between numerous systems; of which XML provides the capabilities to do so.

3.1.2 XML Schema

XML Schema is one of several XML schema languages. Currently it is an ongoing endeavour of W3C to replace XML DTD (Document Type Definition), a common XML schema language with XML Schema [46]. This is mainly due to XML Schema being considered to be more expressive and more easily used by a wider variety of applications [46]. XML schemas can also be used to define the content and structure of XML documents” [47]. An example of an XML schema, used to store data recorded from a device, is presented in Figure 3:

![XML Schema Example](image)

**Figure 3.** An XML Schema Example

3.2 Data Storage

Data generated within a smart home environment will be stored electronically and accessed by some form of agent. A common means of storing data is a database, a central location for related records and files [48].

XML documents can be considered as containing either, document-centric or data-centric data. Typically, data-centric data is updateable and stored in a well-defined structure [49]. However, in comparison, document-centric data is typically irregular in content and size [49]. As XML continues to grow in popularity an increasing number of systems are required to manage both types of data. Therefore, XML storage systems must be capable of efficiently accommodating both data types [49]. Typically relational databases are more appropriate for data-centric data whilst document management and content management systems are more appropriate when dealing with document-centric data [49]. In both instances XML documents must be mapped directly to the data stored in the database, to ensure data can be moved between XML documents and the database [49]. Finally, to ensure the integrity of a database, all data must be stored in a consistent format and in a reliable and efficient way [50].
3.2.1 Native XML Database v XML-Enabled Database

The use of a database allows for a more sophisticated means of storing and retrieving data [51]. There are 2 different methods that can be used to store an XML document; either using an XML Enabled Database or a Native XML Database.

XML-Enabled Databases

XML-Enabled Database typically stores documents consisting of data-centric data, by mapping the XML document’s data model to a database model and converting the XML data into the database [51]. There are a number of commercial XML-Enabled database products, such as Microsoft SQL Server 2000 and Oracle’s Oracle8i. Within an XML-Enabled database, each XML document will be stored as one single column [49]. Due this method of data storage problems can arise when attempts are made to index or search a document [49].

In XML-Enabled databases, each XML document schema will correspond to an XML schema; with the XML external to the database and invisible inside the database [51].

Native XML Databases

Native XML Databases are considered to be a more efficient means of storing document-centric XML documents [51]. They are designed specifically to store XML documents [52] and therefore do not need to be translated into an object based or relational database structure [49].

An XML schema within Native XML databases record rules regarding indexing and the storage of data, as well as any information that needs to be conveyed to the Native XML database engine [49]. Finally, Native XML databases are considered to be particularly suitable for storing "irregular, deeply hierarchical, and recursive data" [51].

3.3 Ontologies

An ontology conceptualises a domain and will represent relationships and concepts in a format that is machine-readable whilst remaining user-understandable [53]. They are becoming increasingly present in information systems; within diverse fields such as medicine and e-commerce [54]. Ontologies can be used to analyse information to enhance it [55]. Ontologies play a significant role when exchanging information by providing a common understanding of a domain [56].

Within the healthcare sector each individual patients information can extend over a range of departments that may not necessarily interoperate; resulting in difficulties arising when a doctor requires a full medical history of a patient. Medicine is a domain rich in knowledge defined through standards; which would be greatly assisted through the development of ontologies [57]. It is not
always necessary for each institute to develop entirely new ontologies; as concepts from existing standards within the healthcare domain [57] such as HL7 [58] can be used.

3.4 Similar Data Storage and Formatting Projects

The section below will discuss current data storage and formatting projects within the scientific domain; in particular projects within the healthcare and smart home research domain.

**NASCArray**

NASCArrays is a repository available to researchers within the Arabidopsis domain, since 2002 to store microarray data generated by NASC’s transcriptomics service [59]. Users perform their own experiments before sending the data to NASC for further processing and data preparation [59]. All data that is produced by NASC is made public via the NASCArrays database, where the data can be downloaded as a comma separated values (CSV) file [59].

3.4.1 Approaches in Healthcare

A sample of data storage and formatting projects within the healthcare domain will be discussed within this section.

**ecgML**

Electrocardiogram (ECG) data exchange and comparability can be limited between researchers as there are a range of formats and platforms that can be used to generate and store the data [60]. Therefore, ecgML has been developed using XML technology to create a solution that is format and system independent or representing ECG data [60].

In addition to the development of an ECG data protocol a suite of tools are also being created to be used by researchers; including XML-based an ECG parser, an ECG record generator, and an ECG viewer [60].

**ECG-BSPM**

The Body Surface Potential Map (BSPM) is a method for displaying and recording electrocardiographic data [61]. ECG-BSPM was developed as a format for the storage of BSPM datasets, using XML technology. A large number of BSPM electrode configurations exist, and therefore, the BSPM format was designed to be versatile and support the various BSPM configurations in use today [61].

**ECGAware**

ECGAware is an XML-based format for the storage of ECG data. The format aims to be application and platform independent whilst incorporating elements arising from the current growth in telemonitoring [62].
**OpenECG, SCP-ECG**

The aim of the OpenECG project was to promote interoperability in the area of electrocardiography. The Standard Communication Protocol for Computer-Assisted Electrocardiogram (SCP-ECG) has been developed through the OpenECG project and is a European standard for the digital storage and exchange of ECG data [63]. It specifies the format, as well as the messaging procedure for the exchange of ECG data between computer systems and electrocardiograms [63]. The OpenECG portal is also available providing a range of services designed to assist users wishing to adopt the standard [63].

**DICOM**

The Digital Imaging and Communications in Medicine (DICOM) standard specifies the file format for biomedical images [64]. The standard is being continuously extended modularly in order to support new technologies and applications [64].

**FDADF**

After performing a survey on ECG standards the Food and Drug Administrator (FDA) chose XML as the technology for data representation and as a result in 2002, they produced the FDA XML Data Format (FDADF); which is an effort to standardise ECG data for all stakeholders [62].

**HL7**

HL7 [58] is a non-profit, standards developing organisation dedicated to developing standards for the exchange, management and integration of electronic healthcare information using XML to enable the interoperability of healthcare documents [58]. The need for such a standard stems from the desire to unlock the considerable clinical content stored in free text clinical notes, and to enable the pooling of content from numerous documents created on various systems [65].

**An XML Based Multimedia Data Acquisition and Retrieval with Wearable Computers**

Ohmori et al. [66] have presented a inquiry and data retrieval system for wearable computers that is XML based [37]. Multimedia data recorded in mobile environments will be stored in an XML native database adhering to XML formats [66]. A means of allowing researchers to search results in an appropriate form will also be available [66].

Finally, Fengguang et al. [47] discuss the use of XML as a way of storing data in a simple and rapid way to enable the integration of data stored from various data sources. While Catley and Frize [67] present the potential application of a standardized medical database and integrated medical applications using XML along with the Unified Medical Language Systems (UMLS). Currently most medical databases are institution dependent and as such have a heterogeneous nature. It is hoped that
through the use of XML, data sharing between institutions could be enhanced and challenges posed by non-existent medical database standardization could be solved.

3.4.2 Approaches related to Smart Homes

The section below will discuss data storage and formatting projects related to the data generated within smart home environments.

XML Modelisation of Smart Home Environment

People with disabilities often rely on assistive technologies, such as electronic wheelchairs, to assist people with limited movement, robots to move objects within the home and computers to improve communication [68]. Consequently, the user will be presented with several heterogeneous systems, each with different user interfaces and functionality [68]. Therefore, Feki et al. [68] have developed models for the development of XML based objects to support interactions between systems and devices within an environment [37]. Manufacturers are invited to hide their product specification and represent them in a unified way, making it possible to present an appropriate design to those with disabilities [68].

A software configuration tool, called ECS (Environment Configuration System) has been developed to allow a non expert to configure any input device within a smart home environment [68]. Equipment will be described using XML to automatically generate all available functionality's and display them in an interactive graphical user interface [68]. To date, ECS is compatible with most home equipment, generating an XML object as a standard output which can be easily downloaded within the system [68].

AALLIANCE

The AALIANCE project is focused on developing Ambient Assisted Living (AAL) solutions in the areas of ageing at work, ageing in society and ageing at home [69]. AALIANCE aims to develop a common strategic vision for short to long term research and development approaches in the field of AAL [69]. They have also been particularly successful in defining standardisation requirements within this area [69]. AALIANCE considers the availability of standards within the AAL domain to be inadequate, therefore, they have suggested steps [70] that should be taken to resolve this issue including:

- Increase the awareness of existing standards within the AAL research domain.
- Promote the advantages of application and services collaborations based on standards.
- Develop and promote a reference model that provides guidance to developers.
- Develop and encourage the use of design guidelines and a certification process for AAL services and products.
• Move from syntactical to semantic and process interoperability levels.
• Encourage research to develop standards for semantic and process interoperability levels.

**OASIS**

Open Architecture for Accessible Services Integration and Standardisation (OASIS) has an overall goal of revolutionising the interoperability and quality of services, available to older people to assist them as they perform their daily activities [71]. The OASIS platform is open and modular and is considered to be easy to use to enable the interoperability of content among ontologies and systems in all applications related to the elderly [71]. A set of tools for content/services connection and management, user interface creation and adoption are included within the platform [71].

The overall aim of OASIS is 'direct re-usability of information is to be provided across heterogeneous services and devices' [72]; through the combination of 3 component technologies [72] listed below:

• The use of restricted ontologies to access web-based services.
• The use of formal foundational ontologies to capture device and services semantics.
• The use of standards and methods developed for structured software engineering.

**SensorML**

SensorML is an Open Geospatial Consortium (OGC) approved international, open technical specification; providing standard models and XML encoding to describe sensors and measurement processes [73]. SensorML defines processes as "inputs, outputs, parameters, and methods, as well as provide relevant metadata" [73]. In its simplest application, SensorML provides a standard digital means of creating sensor component and systems specification sheets [73].

**Sensory Dataset Description Language (SDDL) Specification**

SDDL is an XML-encoded description language for sensor data generated within pervasive spaces, and is proposed as a standard for characterising sensor datasets to improve interoperability and the ability to share datasets among members of the research community [74]. The description language will specify collective information including sensors/actuators, dataset parameters and sensor events; however it does not specify any physical properties of the sensors/actuators [74].

Although the applications and studies discussed above have been successful in their own application domains, the issue of addressing the need for an open format to support the exchange and storage and exchange of data generated within intelligent environments has not yet been addressed [37].
3.4.3 Key Research Projects

Below I have discussed 2 projects that have particularly inspired this study:

PhysioNet is an online source to vast amounts of open-source software and physiological signals; with the intended purpose of stimulating current research and investigations in the area of biomedical and physiological signals [75]. PhysioNet offers 3 components, PhysioBank, PhysioToolkit and PhysioNet. PhysioBank offers a large archive of digital recordings of Physiologic signals conforming to any Wave Form Database (WFDB) compatible format while PhysioToolkit is a large library offering access to open source software that can be used by researchers to process and analysis physiological signals and finally, the PhysioNet website enables free and open distribution of open-source software and physiological signals within the research domain [75].

During a workshop within CHI’09 [76] the issues surrounding the development of shared home behaviour datasets were discussed to advance home-computer interaction and ubiquitous computing research [77]. Intille [78] has continued this research and outlines a problem within an area he is currently working, ‘Portable In-Home Data Collection’; whereby work related to context-awareness within the home is being limited due to the lack of large datasets available to researchers to test their developments and discoveries [78]. Intille has proposed the development of a community resource containing six datasets consisting of high quality, synchronised data streams recorded over a four month period from most sensor types currently being used within smart home environments [78]. This resource will enable researchers to focus on development and testing without being stalled by the need for data collection.

4.0 Contribution

The primary contribution from this study will be the continued development of an online repository and suite of tools that will be available to members of the research domain in order to assist them as they perform their research. An end-to-end system will be produced allowing researchers within this domain to design their experiment, upload and store the data generated through their experiment and finally visualise the data in order to recognise any changes to a patients behaviour. A gap has been identified within the current research domain that could be filled with the development of such an online application. The suite of tools will be developed within the Visual Studio development environment using C#, which aids the rapid development of user interfaces through support for drag-and-drop user interface development, whilst facilitating customisation of user interface components through custom C# code.

To date there does not exist a standard within the research community specifying how data generated within a smart home environment should be stored. Due to this there is a growing need along with an opportunity to develop a standard format for its storage and exchange.
Initially the work will focus on the development of an XML-based format; extending upon the concepts of homeML [37], for storing data generated by a range of devices and services both within the home environment and also outside of the home.

homeML was originally proposed as a means of solving the issues caused by the heterogeneous nature of data recorded within intelligent environments through the creation of an XML-based format for data generated within intelligent environments [37]. It is the initial aim of this research to create a solution to the problems caused by the heterogeneous nature of data generated within a smart home environment, extending upon the previously proposed solution of homeML which aimed to develop a standard format for the exchange and storage of data. XML has been chosen as the mark-up language of choice as it is application and platform independent, which will be greatly beneficial in regard to interoperability and sharing between users.

The contribution to knowledge will be realised through answering the following questions:

1. **Can homeML be extended to incorporate smart environments outside of the home?**

   Currently homeML version 1.0 is designed to store data generated within a smart home environment, however, as technology continues to evolve data can be generated outside of the home environment. Therefore, homeML version 1.0 should be extended and validated to incorporate this functionality.

2. **Can a suite of tools be developed and a central repository created to assist researchers within this domain?**

   Through the creation of a centralised data repository and complementary suite of tools, researchers within this domain will have access to an efficient method for designing experiments and recording data generated within smart environments; whilst adhering to a standard format.

3. **Is homeML accepted within the research domain as a standard format for the exchange and storage of data generated within smart environments?**

   It is important that the final homeML schema is critically reviewed and accepted by a number of research institutes; to ensure homeML becomes accepted as the standard format for the exchange and storage of data generated within intelligent environments.

4. **Can sharing decision support rules generated within intelligent environments offer decreased deployment time and increased reliability?**

   HomeRuleML [79] is a proposed model, for the representation of decision support rules for intelligent environments. With the aim of providing researchers within this domain with a set of open rules, extending the current work further [79].
The following sections, Work to Date and Future Work will discuss the work that has been completed or will be undertaken in order to achieve the 4 identified research questions above. A work schedule has also been agreed that can be viewed in Appendix E.

5.0 Work to Date

To date development has been an iterative process shown in Figure 3. The work completed is presented in the following section.

![Flow Chart Depicting the Iterative Development Process](image)

**Figure 4.** Flow Chart Depicting the Iterative Development Process.

**Review and validate the homeML concept.**

homeML [37] was a solution proposed to address the problems caused by the heterogeneous nature of data generated within smart home environments. Initially, homeML was critically reviewed to ensure the rational and accuracy of the concept. Thorough research was undertaken to ensure no other standard had been developed since the time of the original homeML proposal (refer to Section 3). Once confirmed the next stage of the review process involved accessing a number of open-source datasets from various research institutes in order to compare the data structure used by each and the ability of homeML to store this data. The detailed results of this review can be to found in Appendix A.

The review process identified a few inaccuracies in the original schema. Table 5 documents the changes and rational for each adjustment. As a result of this review process homeML version 1.0, which can be viewed in Appendix B, was altered to incorporate the adjustments and moved to homeML version 1.1.
<table>
<thead>
<tr>
<th>Attribute</th>
<th>Alteration</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Room</td>
<td>'RoomDescription' is now optional rather than required.</td>
<td>Only one institute recorded room description and so it was decided that this should be optional data rather than compulsory.</td>
</tr>
<tr>
<td>Room</td>
<td>'Floor' is now optional rather than required.</td>
<td>No dataset included floor level and so it was decided that this should be optional data rather than compulsory.</td>
</tr>
<tr>
<td>Room</td>
<td>'Units' is now optional rather than required.</td>
<td>No dataset included device units and so it was decided that this should be optional data rather than compulsory.</td>
</tr>
<tr>
<td>Room</td>
<td>'Data' was added to 'RealTimeInformation' - previously only 'SampleRate', 'StartTime' and 'EndTime' where recorded.</td>
<td>Although SampleRate, StartTime and EndTime were recorded, the data actually being produced by the device had no place to be stored and therefore the 'Data' tag was required.</td>
</tr>
<tr>
<td>Ambulatory Device</td>
<td>This attribute is now considered as optional.</td>
<td>Although, a patient is living within a smart home environment they will not necessarily be required to wear an ambulatory device.</td>
</tr>
<tr>
<td>Ambulatory Device</td>
<td>'Units' is now optional rather than required.</td>
<td>No dataset included device units and so it was decided that this should be optional data rather than compulsory.</td>
</tr>
<tr>
<td>Ambulatory Device</td>
<td>'Data' was added to 'RealTimeInformation' - previously only 'SampleRate', 'StartTime' and 'EndTime' where recorded.</td>
<td>Although SampleRate, StartTime and EndTime were recorded, the data actually being produced by the device had no place to be stored and therefore the 'Data' tag was required.</td>
</tr>
</tbody>
</table>

*Table 5. The alterations and corresponding rational made to homeML version 1.0*

*Work with open-source data sets in order to establish data formatting requirements*

A number of open-source datasets are available online from various institutes, conducting research in the area of intelligent environments, [http://boxlab.wikispaces.com/List+of+Home+Datasets](http://boxlab.wikispaces.com/List+of+Home+Datasets). The datasets used within this study included Massachusetts Institute of Technology [80], University of Amsterdam [81], University of Texas at Arlington [82], Washington State University [82] as well as
datasets produced within the University of Ulster at Jordanstown [83]. The datasets are each recorded in an individual format and stored as various file types including Microsoft CSV files and IN files.

As previously mentioned, during the review process of homeML version 1.0 these datasets were accessed in order to compare the data structure used by each to the homeML version 1.0 data schema. In addition, the datasets were also useful in establishing data formatting requirements, enabling homeML version 1.1 schema to be extended further.

Extension of homeML 1.0 incorporating environments external to the home

As technology continues to evolve within this domain it is now possible to monitor a patient outside of their smart home environment; with a significant push in the direction of ambient assisted living (AAL). The data generated through the patients interactions with their environment will differ as a patient moves between locations. Upon closer inspection, it was decided that the homeML version 1.1 schema had limited use in regard to the storage of data generated within environments external to the home. Therefore, homeML version 1.1 was extended further to incorporate the need to consider a patient's location when recording data. The hierarchical data trees can be viewed in Figure 5. Tables 6 and 7 provide additional information to support the design of the data tree.

As depicted by the hierarchical representation a device can be categorised as being either a location or a mobile device. A location device is a device restricted to one location, whether it is integrated into the environment or not. PIR movement sensors that are restricted to one location would be an example of this, or an ECG device, used to measure a patient's heart rate at regular intervals, that is not integrated into an environment but is going to be used in one location at a specific point in time. A mobile device, however, is not location restricted and can easily move between locations. An accelerometer would be an example of this type of device which is capable of being used continuously as a patient moves between environments.

In addition to data being generated through a patient’s interactions with devices and services integrated into an environments, such as door sensors and PIR sensors; data can also be recorded using mobile devices to record data as a patient moves between environments, such as a mobile phone recording a patient’s location using GPS tracking.
**Figure 5.** Hierarchical representation of homeML version 2.0 as a series of Tree diagrams.

<table>
<thead>
<tr>
<th>HomeML</th>
<th>The root element for XML based smart home data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Element/Attribute</td>
<td>Description</td>
</tr>
<tr>
<td>InhabitantDetails</td>
<td>Personal Information of inhabitant</td>
</tr>
<tr>
<td>Location</td>
<td>Definition of the location and devices within a location</td>
</tr>
<tr>
<td>MobileDevice</td>
<td>Definition of mobile device</td>
</tr>
</tbody>
</table>

**Table 6.** Description of elements within homeML version 2.0, *Figure. 5.*

<table>
<thead>
<tr>
<th>Location</th>
<th>The Location Element for XML based smart home data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Element/Attribute</td>
<td>Description</td>
</tr>
<tr>
<td>LocationDescription</td>
<td>General description of the location</td>
</tr>
<tr>
<td>LocationClass</td>
<td>Details the type of location, e.g. outside, smart home</td>
</tr>
<tr>
<td>LocationDevice</td>
<td>Definition of location device</td>
</tr>
</tbody>
</table>

**Table 7.** Description of Location element within homeML version 2.0, *Figure. 5.*

**Evaluation of homeML version 2.0**

A similar process was involved when reviewing and validating homeML version 2.0 as was involved when reviewing version 1.0. Again, a number of open source data sets from various institutes were accessed and used as a means of comparing their current structure to homeML version 2.0, including Massachusetts Institute of Technology [80], University of Amsterdam [81], University of Texas at Arlington [82], Washington State University [82] as well as datasets produced within the University
of Ulster at Jordanstown [83]. In addition a number of scenarios were performed, details of which can be found in Appendix C in order to generate new data that could also be input into the XML schema in order to assess whether the structure was compatible with a range of data sources and structures. Based on this review process, which can be viewed within Appendix D there were no changes identified to the defined schema. Figure 6 provides an example of data stored according to the homeML version 2.0 schema.

```xml
<Location LocationID="1">
  <LocationDescription>LocationName</LocationDescription>
  <LocationClass>Indoor/LocationClass</LocationClass>
  <LocationDevice ID="00000100">
    <DeviceDescription>DeviceDescription</DeviceDescription>
    <DeviceLocation>
      <XPos>YPos</XPos>
      <YPos>ZPos</YPos>
    </DeviceLocation>
    <DeviceType>Type</DeviceType>
    <DeviceID>DeviceID</DeviceID>
    <RealTimeInformation>
      <SampleRate><SampleRate></SampleRate>
      <StartTime><EndTime></EndTime>
      <Date></Date>
    </RealTimeInformation>
    <Event EventID="1">
      <TimeStamp>2010-02-08 11:25:51</TimeStamp>
      <Data></Data>
    </Event>
  </LocationDevice>
</Location>
```

Figure 6. Example of data according to the homeML version 2.0 schema.

**homeML version 2.0 XML Creator**

The first stage of development has involved creating a user interface, in C# that will allow a user to create an XML schema, tailored to the data they wish to store, as depicted in Figure 7. The user is able to select the data options they would like to store by ticking the check boxes corresponding to each tag. The essential data, such as InhabitantDetails and Location, however, are not optional and will always be created in the XML schema.

![homeML XML Creator Interface](image)

Figure 7. homeML XML Creator Interface
Once, the user clicks on the 'Create XML Schema' button, the XML schema will be created according to the specifications of the user, as shown in Figure 8. This tool will provide researchers within this domain with the opportunity to design their experiments and the format the data should adhere to, prior to their experiment commencing.

![Figure 8. homeML XML Schema Example](image)

### 6.0 Further Work

The following section will discuss work to be completed. As well as this a proposed thesis outline has also been discussed and can be viewed in Appendix F.

**The development of a suite of tools that will be available to members of the research community.**

An online interface will be developed in C# providing a suite of tools that will be available to members of the research community. This suite of tools will allow researchers to perform a number of tasks including:

- Upload data generated within a smart home environment conforming to the homeML version 2.0 standard.
- Define a number of rules regarding a patient's ADL (activities of daily living) that can be used to diagnose and monitor a patient's condition.
- Allow researchers to create dynamic profiles for patients, allowing behaviour changes to be recognised according to the data gathered from their environment; through the implementation of a proposed method, HomeRuleML [79].

The first step will be to develop an interface that will enable researchers to upload their data gathered from various sources to one central repository. It is hoped that eventually an end-to-end system will be developed and used by researchers within this domain, similar to PhysioNet [75]. Whereby, a user will have access to a database that will allow them to design experiments, providing details of an environment and the devices within it. Finally, there will be a means of visualising the data the user has uploaded and converted to the homeML version 2.0 schema, an example of the visualising it is hoped will be achieved can be found in Figure 9.
Figure 9. A graph illustrating the number of times the main door sensors within 16J26 and 16J27 were activated each hour over a period of 24 hours.

The development of a means to store the data generated within smart home environments within a central repository.

A centralised data repository in its simplest form is a data warehouse capable of storing large amounts of data from multiple sources [84]. A method will also be developed allowing a researcher to store data generated within intelligent environments within a central repository adhering to a specified format.

As the data generated within this domain is of a heterogeneous nature it is essential that it is organised and stored within a central repository according to a preset specification. Conforming to a format specification can make data easier to exchange and store, as well as more comparable and meaningful. The raw data generated within a smart environment does not require re-engineering and altering but should be organised and saved according to a pre-specified format and appropriate structure. Data standardisation is involved during this process and will include transforming the data to an agreed form and in doing so address various formatting issues and improve data interoperability in addition to user understanding of the data itself [85].

The data generated within intelligent environments will be have a regular structure and have little or no mixed content [52] and will produce XML documents that can be classified as data-centric. It will therefore be more suitable and efficient to use an XML-enabled database as the centralised data repository. Whereby, the data will appear in a regular order and will be stored in a database, while the XML will exchange and publish the data[50]. One of the major benefits of using this type of database is that the XML data tags do not need to be recorded within the database, therefore reducing the storage capacity of a document and improving transmission and processing performance [49].

Through the development of a centralised data repository and associated data format; the exchange, storage and analysis will be dramatically improved.

The promotion of homeML as a standard for the exchange and storage of sensory within the research domain of intelligent environments.
To achieve the acceptance of homeML version 2.0 as a standard within the research domain, a number of researchers from various institutes will be asked to validate the XML schema; in order to obtain feedback regarding the structure and any suggestions they may have.

A research group from Singapore have also proposed to submit homeML as a case study for consideration by the Technology Working Group of the Continua Health Alliance for the definition of new concepts for modelling in Smart Homes. They have been made aware of our interest in working with them on this and they have agreed to proceed in a collaborative submission. Although, at this stage we are unsure of what this will involve we are obviously very excited as this may result in homeML being widely promoted.

Finally, submissions have already been made to ITAB 2010 (Appendix H) and NIBES'10 (Appendix I) documenting the work performed to date as well as future work plans.

7.0 Conclusion

Due to the dramatic increase in the global population the current approaches to health and social care provision are being improved and directed in the area of home based support and healthcare provision. This has resulted in a wide range of solutions and services becoming available to monitor and support individuals within and outside their homes in addition to many types of assistive technologies that can be used to augment multiple activities of a person’s day-to-day life; some of which have been discussed within this report.

Within the report the problem area has been discussed and identified as the abundance of heterogeneous data being generated and subsequently stored in various formats; limiting the opportunities to exchange, share and compare data. A number of standard data formats exist within the healthcare domain and have been discussed within this report; however a gap in the current literature has identified that there does not exist a standard addressing the exchange and storage of data within intelligent environments.

It is the aim of this study to develop such a standard, extending upon a proposed solution homeML [16] which aimed to develop a standard format for the storage and exchange of data, providing a solution to the problems caused by the heterogeneous nature of the data produced in and outside of a smart home environment.
8.0 References


"Ambient Assisted Living (AAL) Joint Programme: Objectives," [Online]. Available: 


Appendix A: homeML version 1.0 review

Results showing the review of homeML version 1.0 as well as the properties of each dataset used:

<table>
<thead>
<tr>
<th>home ML</th>
<th>Date of Recording</th>
<th>Room Description</th>
<th>Room Param</th>
<th>Floor</th>
<th>Device (deviceID)</th>
<th>Description</th>
<th>Device Location (X,Y,Z)</th>
<th>Device Type</th>
<th>Units</th>
<th>Real Time Information (runID)</th>
<th>Event (eventID)</th>
<th>Time Stamp</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>UUJ</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y (sensorID)</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>UA</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y (sensorID)</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>UTA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MavLab</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y (sensorID)</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>MavPad</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y (level)</td>
<td>Y (number)</td>
<td>Y (info)</td>
<td>Y (zone)</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>WSU</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y (ID)</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>MIT</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y (sensor_ID)</td>
<td>N</td>
<td>Y (zone)</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
</tbody>
</table>
## Dataset Properties:

<table>
<thead>
<tr>
<th>Institute</th>
<th>Experiment</th>
<th>File Name</th>
<th>File Type</th>
<th>Size</th>
<th>Date</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>UUJ</strong></td>
<td>GPS Test Data</td>
<td>gpsData.text</td>
<td>Text</td>
<td>141 KB</td>
<td>08/03/10</td>
<td>[88]</td>
</tr>
<tr>
<td></td>
<td>Accelerometer Test Data</td>
<td>accData.text</td>
<td>Text</td>
<td>189 KB</td>
<td>08/03/10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kitchen Sensor Recordings</td>
<td>Mark0958.csv</td>
<td>CSV</td>
<td>1 KB</td>
<td>22/10/07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trackstick Test Data</td>
<td>Test.tsf</td>
<td>TSF</td>
<td>11 KB</td>
<td>25/02/10</td>
<td></td>
</tr>
<tr>
<td><strong>UA</strong></td>
<td>Activity Recognition from Binary Sensors</td>
<td>kasterenActData.txt</td>
<td>Text</td>
<td>12 KB</td>
<td>02/07/08</td>
<td>[74]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>kasterenSenseData.txt</td>
<td>Text</td>
<td>62 KB</td>
<td>02/07/08</td>
<td></td>
</tr>
<tr>
<td><strong>UTA</strong></td>
<td>MavPad 2005 Sensor Data</td>
<td>1-3-2005-RAW.data</td>
<td>DATA</td>
<td>5,433 KB</td>
<td>15/05/08</td>
<td>[75]</td>
</tr>
<tr>
<td></td>
<td>MavLab Sensor Data</td>
<td>March.in</td>
<td>IN</td>
<td>47 KB</td>
<td>15/05/08</td>
<td></td>
</tr>
<tr>
<td><strong>WSU</strong></td>
<td>WSU WorkPlace Testbed</td>
<td>Ailab_Spring2008.txt</td>
<td>Text</td>
<td>26,926 KB</td>
<td>24/02/09</td>
<td>[75]</td>
</tr>
<tr>
<td></td>
<td>Apartment Testbed, ADL Multiresident Activities</td>
<td>P01.txt</td>
<td>Text</td>
<td>27 KB</td>
<td>24/02/09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Apartment Testbed, Normal ADL activities</td>
<td>P01.t1</td>
<td>T1</td>
<td>3 KB</td>
<td>22/05/08</td>
<td></td>
</tr>
<tr>
<td>MIT</td>
<td>Activity Recognition in the Home Setting Using Simple and Ubiquitous Sensors</td>
<td>activities_data.csv</td>
<td>CSV</td>
<td>97 KB</td>
<td>12/10/04</td>
<td>[73]</td>
</tr>
<tr>
<td>-----------</td>
<td>------------------------------------------------------------------------------</td>
<td>----------------------</td>
<td>-----</td>
<td>-------</td>
<td>----------</td>
<td>------</td>
</tr>
<tr>
<td>UUJ</td>
<td>University of Ulster, Jordanstown</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UA</td>
<td>University of Amsterdam</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UTA</td>
<td>University of Texas at Arlington</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WSU</td>
<td>Washington State University</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MIT</td>
<td>Massachusetts Institute of Technology</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix B

homeML version 1.0 hierarchical tree diagram:

[Diagram of homeML version 1.0 structure]

KEY
+ Must exist at least once
* Can exist any number of times
? Optional
Nodes without any sign must exist once
Appendix C: Scenarios

Scenarios used as a means of generating new data used to verify homeML version 2.0:

Scenario 1: In-Home

*Purpose:* To track and measure a patient's movement within their home environment.

*Devices:* Door sensors, PIR sensors and an accelerometer (worn by patient).

*Test Case:* Patient moves from living room to kitchen (can be performed in either J26 or J27)

- Living Room - PIR (on)
- Living Room - Door (open)
- Kitchen - PIR (on)

  - At the same time the accelerometer is continuously measuring the distance and the speed at which the patient moves.

Scenario 2: In-Home

*Purpose:* To monitor a patient's heart rate during regular intervals as well as track and measure a patient's movement within their home environment.

*Device:* ECG Monitor, Door sensors, PIR sensors and an accelerometer (worn by patient).

*Test Case:* To take ECG readings from the patient; whilst also measuring a patient's movement from living room to kitchen (can be performed in either J26 or J27).

- Living Room - PIR (on)
- Living Room - Door (open)
- Kitchen - PIR (on)

  - At the same time the accelerometer is continuously measuring the distance and the speed at which the patient moves and the ECG monitor is used to take regular readings from the patient.
Scenario 3: Out-of-Home

*Purpose:* To continue to track and measure a patient's movement if they leave their home environment.

*Devices:* Door Sensors, PIR sensors, accelerometer (worn by patient) and mobile device (built in GPS).

*Test Case:* Patient leaves home environment

- Hallway - PIR (on)
- Front Door - open
- Front - Door (closed)
- GPS automatically starts

  - Again the accelerometer continues to measure the distance and speed at which the patient is moving, whilst tracking their location.
Appendix D: homeML version 2.0 review

Results showing the review of homeML version 2.0:

<table>
<thead>
<tr>
<th>Home ML</th>
<th>InhabitantDetails</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>InhabitantID</td>
<td>Name</td>
</tr>
<tr>
<td>UUJ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UTA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MavLab</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MavPad</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WSU</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MIT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home ML</td>
<td>MobileDevice</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>--------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DeviceType</td>
<td>Units</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Run ID</td>
<td>Sample Rate</td>
</tr>
<tr>
<td>UUJ</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>UA</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>UTA</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>MavLab</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>MavPad</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>WSU</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>MIT</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Home</td>
<td>ML</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>----</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DeviceType</td>
<td>Units</td>
<td>DevicePlacement</td>
</tr>
<tr>
<td>RunID</td>
<td>SampleRate</td>
<td>StartTime</td>
</tr>
<tr>
<td>UUJ</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>UA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UTA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MavLab</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MavPad</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WSU</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MIT</td>
<td>Y</td>
<td></td>
</tr>
</tbody>
</table>

|                  | University of Ulster, Jordanstown |
| UUJ              | University of Amsterdam          |
| UA               | University of Texas at Arlington |
| UTA              | Washington State University      |
| WSU              | Massachusetts Institute of Technology |
## Appendix E - Proposed Work Schedule

<table>
<thead>
<tr>
<th>Task Name</th>
<th>Duration</th>
<th>Start</th>
<th>Finish</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Research Training</td>
<td>717 days</td>
<td>Thu 01/10/09</td>
<td>Fri 25/06/12</td>
</tr>
<tr>
<td>2. Literature Review</td>
<td>195 days</td>
<td>Thu 01/10/09</td>
<td>Wed 30/06/10</td>
</tr>
<tr>
<td>3. Identify Problem Area</td>
<td>1 min</td>
<td>Thu 01/10/09</td>
<td>Wed 20/10/09</td>
</tr>
<tr>
<td>4. Identify Similar Projects</td>
<td>3 mos</td>
<td>Mon 02/11/09</td>
<td>Fri 22/01/10</td>
</tr>
<tr>
<td>5. Investigate Technology Tools</td>
<td>1 min</td>
<td>Mon 22/01/10</td>
<td>Fri 18/02/10</td>
</tr>
<tr>
<td>6. Identify Research Gap</td>
<td>1.5 mos</td>
<td>Mon 22/01/10</td>
<td>Fri 14/02/10</td>
</tr>
<tr>
<td>7. Identify Contribution to Knowledge</td>
<td>1.5 mos</td>
<td>Mon 22/01/10</td>
<td>Fri 14/02/10</td>
</tr>
<tr>
<td>8. Prepare Confirmation Report</td>
<td>2 mos</td>
<td>Mon 05/05/10</td>
<td>Fri 25/06/10</td>
</tr>
<tr>
<td>9. Extend homRUL to incorporate smart environments outside of</td>
<td>4 mos</td>
<td>Mon 01/05/10</td>
<td>Fri 21/05/10</td>
</tr>
<tr>
<td>10. Review and Validate the homRUL concept</td>
<td>14 days</td>
<td>Mon 01/05/10</td>
<td>Thu 18/02/10</td>
</tr>
<tr>
<td>11. Work with open source datasets to identify data formatting requirements</td>
<td>1 mos</td>
<td>Fri 19/02/10</td>
<td>Thu 18/03/10</td>
</tr>
<tr>
<td>12. Extend homRUL, version 1.0</td>
<td>1.5 mos</td>
<td>Fri 19/02/10</td>
<td>Thu 25/04/10</td>
</tr>
<tr>
<td>13. Evaluate and Test homRUL, version 2.0</td>
<td>16 days</td>
<td>Fri 19/02/10</td>
<td>Fri 21/05/10</td>
</tr>
<tr>
<td>14. Dissertation - Conference</td>
<td>1 mos</td>
<td>Mon 03/05/10</td>
<td>Fri 28/05/10</td>
</tr>
<tr>
<td>15. Create suite of tools and online repository</td>
<td>259 days</td>
<td>Thu 01/07/10</td>
<td>Tue 26/06/11</td>
</tr>
<tr>
<td>16. Literature Review</td>
<td>1.5 mos</td>
<td>Thu 01/07/10</td>
<td>Wed 11/08/10</td>
</tr>
<tr>
<td>17. Soldier Development</td>
<td>131 days</td>
<td>Mon 02/09/10</td>
<td>Mon 31/01/11</td>
</tr>
<tr>
<td>18. Soldier Testing and Evaluation</td>
<td>3 mos</td>
<td>Fri 04/02/11</td>
<td>Thu 28/04/11</td>
</tr>
<tr>
<td>19. Dissertation - Journal</td>
<td>1 mos</td>
<td>Wed 01/09/11</td>
<td>Tue 26/06/11</td>
</tr>
<tr>
<td>20. Dissertation - Conference</td>
<td>1 mos</td>
<td>Wed 01/09/11</td>
<td>Tue 26/06/11</td>
</tr>
<tr>
<td>21. Implement and Extend homRUL</td>
<td>5 mos</td>
<td>Mon 03/09/11</td>
<td>Fri 28/09/11</td>
</tr>
<tr>
<td>22. Review and Validate the homRUL concept</td>
<td>14 days</td>
<td>Mon 02/10/11</td>
<td>Thu 20/11/11</td>
</tr>
<tr>
<td>23. Work with open source rules to identify data formatting requirements</td>
<td>1 mos</td>
<td>Mon 24/11/11</td>
<td>Fri 18/02/11</td>
</tr>
<tr>
<td>24. Extend homRUL, version 1.1</td>
<td>1.5 mos</td>
<td>Mon 21/02/11</td>
<td>Fri 18/04/11</td>
</tr>
<tr>
<td>25. Evaluate and Test homRUL, version 2.0</td>
<td>16 days</td>
<td>Mon 04/04/11</td>
<td>Mon 25/04/11</td>
</tr>
<tr>
<td>26. Incorporate homRUL into suite of tools and repository</td>
<td>2 mos</td>
<td>Mon 02/05/11</td>
<td>Fri 14/06/11</td>
</tr>
<tr>
<td>27. Dissertation - Journal</td>
<td>1 mos</td>
<td>Fri 01/07/11</td>
<td>Thu 28/07/11</td>
</tr>
<tr>
<td>28. Dissertation - Conference</td>
<td>1 mos</td>
<td>Mon 01/08/11</td>
<td>Fri 08/08/11</td>
</tr>
<tr>
<td>29. Promote and Gain Acceptance of homRUL tools and repository</td>
<td>60 days</td>
<td>Mon 01/02/12</td>
<td>Fri 06/06/12</td>
</tr>
<tr>
<td>30. Create Blog</td>
<td>14 days</td>
<td>Tue 08/07/10</td>
<td>Fri 12/07/10</td>
</tr>
<tr>
<td>31. Continuously update blog</td>
<td>605 days</td>
<td>Mon 26/07/10</td>
<td>Fri 26/06/12</td>
</tr>
<tr>
<td>32. Thesis Preparation</td>
<td>1 day</td>
<td>Wed 01/07/12</td>
<td>Wed 01/07/12</td>
</tr>
<tr>
<td>33. Review of Literature</td>
<td>66 days</td>
<td>Fri 01/07/11</td>
<td>Fri 06/08/11</td>
</tr>
<tr>
<td>34. Write up thesis</td>
<td>195 days</td>
<td>Mon 03/10/11</td>
<td>Fri 29/06/12</td>
</tr>
</tbody>
</table>
Appendix F - Thesis Outline

Chapter 1: Introduction to the Ageing-Population

- Chronic Diseases
- Ageing-in-Place

Chapter 2: Introduction to Assistive Technologies

- What are Assistive Technologies?
- What are Smart Home Environments?
- What Assistive Devices and Systems are Available to Monitor a Patient?
  - Inside of the Home
  - Outside of the Home
- Current Smart Home Environments

Chapter 3: Introduction to Data Management

- What is Data Management?
- What is Data Formatting?
  - Introduction to XML
  - Introduction to XML Schemas
- What is Data Storage?
  - Introduction to Native XML Databases
  - Introduction to XML Enabled Databases
- What are Ontologies?

Chapter 4: Literature Review

- Introduction to the Ageing Population
- Introduction to Smart Home Environments
- Introduction to Data Management
Chapter 5: homeML
- Overview of homeML
- Data Requirement assessment
  - Method of Eliciting Data Requirements
  - Review of homeML version 1.0
- homeML version 2.0
  - Review of homeML version 2.0

Chapter 6: Application Development and Design
- Application Overview
- Techniques used to Implement the Application
- Evaluation of Application
  - Evaluation Results
  - Evaluation of User Experience

Chapter 7: homeML Validation
- User Assessment
  - Method of Eliciting User Feedback
  - Questionnaire Design
  - Prototype Testing

Chapter 8: HomeRuleML
- Overview of HomeRuleML
- Implementation of HomeRuleML
- Evaluation of HomeRuleML
  - Evaluation Results

Chapter 9: Summary and Discussion of Results
- Effectiveness of Study Solution
- End User Evaluation
Chapter 10: Discussion, Conclusion and Scope for Further Work

- Summary of Main Objectives
- Discussion and Conclusion
- Future Work
# Appendix G - Completed Research Training Seminars

Training courses completed:

<table>
<thead>
<tr>
<th>Title</th>
<th>Date</th>
<th>RTCs Awarded</th>
</tr>
</thead>
<tbody>
<tr>
<td>Induction for Postgraduate Doctoral Studies</td>
<td>16-09-2009</td>
<td>6</td>
</tr>
<tr>
<td>My Doctorate - Let's get started!</td>
<td>22-09-2009</td>
<td>6</td>
</tr>
<tr>
<td>Induction to the Career Development Centre for first year PhD Students</td>
<td>23-09-2009</td>
<td>3</td>
</tr>
<tr>
<td>Meeting and Managing your Supervisor</td>
<td>24-09-2009</td>
<td>3</td>
</tr>
<tr>
<td>Induction to Ref Works</td>
<td>25-09-2009</td>
<td>2</td>
</tr>
<tr>
<td>Developing Information Skills for Effective Research (Computing and Engineering)</td>
<td>29-09-2009</td>
<td>3</td>
</tr>
<tr>
<td>Introduction to Learning and Teaching for PG Teaching Assistants and Demonstrators [FSTHE - Compulsory]</td>
<td>19-10-2009</td>
<td>3</td>
</tr>
<tr>
<td>Systematic Literature Searching</td>
<td>18-11-2009</td>
<td>3</td>
</tr>
<tr>
<td>Communication and Presentation Skills A</td>
<td>23-11-2009</td>
<td>3</td>
</tr>
<tr>
<td>Communication and Presentation Skills B</td>
<td>24-11-2009</td>
<td>3</td>
</tr>
<tr>
<td>The Literature Review - Online</td>
<td>30-11-2009</td>
<td>3</td>
</tr>
<tr>
<td>Project Management in the Research Context - Online</td>
<td>30-11-2009</td>
<td>3</td>
</tr>
<tr>
<td>Planning Your Research Career (Sciences) – Online</td>
<td>15-12-2009</td>
<td>3</td>
</tr>
<tr>
<td>Research Methods (Sciences) – Online</td>
<td>15-12-2009</td>
<td>3</td>
</tr>
<tr>
<td>Research Ethics 1: Good Research Practice (Research Governance) - Online</td>
<td>18-12-2009</td>
<td>3</td>
</tr>
<tr>
<td>Intellectual Property Rights in the Research Context - Online</td>
<td>18-12-2009</td>
<td>3</td>
</tr>
<tr>
<td>Title</td>
<td>Date</td>
<td>Credits</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>------------</td>
<td>---------</td>
</tr>
<tr>
<td>Managing your Research Supervisor or Principle Investigator – Online</td>
<td>13-01-2010</td>
<td>3</td>
</tr>
<tr>
<td>Selecting a Conference, Presenting and Networking – Online</td>
<td>13-01-2010</td>
<td>3</td>
</tr>
<tr>
<td>Getting Published (Sciences) – Online</td>
<td>28-01-2010</td>
<td>3</td>
</tr>
<tr>
<td>Research Ethics 2: Working with Human Subjects – Online</td>
<td>28-01-2010</td>
<td>3</td>
</tr>
<tr>
<td>Planning Your Career: Next Steps</td>
<td>20-04-2010</td>
<td>5</td>
</tr>
<tr>
<td>Getting Published in Engineering</td>
<td>21-01-2010</td>
<td>3</td>
</tr>
<tr>
<td>How to be an Effective Researcher - Part 1</td>
<td>22-04-10</td>
<td>6</td>
</tr>
<tr>
<td>How to be an Effective Researcher - Part 2</td>
<td>23-04-10</td>
<td>6</td>
</tr>
<tr>
<td><strong>Total:</strong></td>
<td><strong>85</strong></td>
<td></td>
</tr>
</tbody>
</table>
Appendix H - ITAB Paper

An XML Based Format for the Storage of Data Generated within Smart Home Environments

H.A. McDonald, C.D. Nugent, G. Moore and D.D Finlay

Abstract— homeML is an XML-based format for the exchange and storage of data within smart home environments. It was originally proposed as a means of solving the issues caused by the heterogeneous nature of data recorded within such environments. In our current work we have extended upon the concepts of homeML and have subsequently developed an improved format to support the storage and exchange of sensory data collected from heterogeneous sources both inside and outside of the home environment. An XML based format has been developed within this study and has been successfully tested by storing datasets generated by 5 research institutions.

I. INTRODUCTION

STUDIES have shown that all industrial and many developing economies are experiencing a demographic transition from a predominantly younger population to a much larger proportion of older people [1]. The proportion of the global population with an age greater than 60 is expected to increase from 10.0% in 2000 to 21.8% in 2050 and reach 32.2% by 2010 [2]. In addition to the increase in the proportion of elderly people, the average life expectancy has also increased from 75.7 years in 1990 to 79.7 in 2008 and is expected to continuously increase further [3]. Concurrently, the number of children being born is also predicted to decrease, with the ratio of 15 year olds to 65 year
olds dropping from the current 9:1 to 4:1 by 2050 [4]. As a consequence, there will be fewer young people to assist older adults cope with the challenges of ageing [5].

As people grow older they are likely to experience age-related impairments; with the older population being most frail and experiencing the highest incidence of disease and co-morbidity, in addition to suffering a high level of disability [1]. This will result in a major impact on the economy as the demand for health and social services will continue to rise [6]. Although remaining at home can become challenging for an elderly person, many wish to remain in their own homes for as long as possible; with more than 80% of people 65 and older living independently [1]. Research has also shown that patients discharged from hospital to a nursing home live on average 2 1/2 years less than someone the same age living at home [7].

Due to these significant demographic changes there is an increasing need to re-assess the current approaches to health and social care provision, particularly in the direction of smart home environments. With no other domain being a richer target of research and commercial interest than the intersection of health, technology and the home [1]. Therefore, enhancing the overall quality of life of the elderly population, whilst allowing them to maintain their independence.

II. AGEING-IN-PLACE

As people age they may begin to experience challenges at home, however, it is not unreasonable for these people to wish to remain at home and maintain their independence. Ageing-in-place refers to members of the elderly community continuing to live in their own home environment, typically not in a healthcare environment, using products, services and conveniences which allow them to remain at home as their circumstances change [8].

A. Assistive Technologies

Assistive Technologies can be defined as “any item, piece of equipment, or product system, whether acquired commercially off the shelf, modified or customized, that is used to increase, maintain or improve the functional capabilities of individuals with disabilities” [9]. Increasingly, assistive technologies are being used to support everything from medication management to personal mobility, with smart applications being employed to monitor, manage, and motivate health behaviours both inside and outside of the home 24 hours a day, 365 days a year [10].

B. Smart Home Environments

The term ‘smart home’ refers to residences equipped with technologies that allow the monitoring of its inhabitants with the ability to encourage independence and the maintenance of good health [4]. Due to the demographic changes discussed above, significant effort has been directed
towards research in the area of home based support and health care provision. This has resulted in a wide range of solutions and services becoming available which can be used to both monitor and track individuals within and outside of their homes. As well as this, many types of assistive technologies have also been developed to augment multiple activities of a person’s day-to-day life.

III. DATA HETEROGENEITY

The main issue with the heterogeneity of data generated within a smart home environment is its lack of interoperability. This creates difficulties when data is being stored and exchanged and also when it is being re-used and most importantly processed by a number of services [11].

An increasing number of researchers wish to access and manipulate data generated from smart environments from a multitude of data sources and achieve interoperability between systems capable of both monitoring and assisting inhabitants. On a general level there are a number of XML based standards within the healthcare domain specifying how data may be stored. For example; Wang et al. [12] have developed ecgML, an application and format-independent solution for the representation and exchange of electrocardiogram data and HL7 [13], a non-profit organisation who develop standards for the exchange, management and integration of electronic healthcare information using XML [14].

As yet there does not exist a standard within the research community specifying how data generated within a smart home environment should be stored. Nevertheless, a number of preliminary studies have been conducted [Intille [15], SensorML [16]]. Due to the issues caused by the heterogeneous nature of the data there is a growing need to develop such a standard. This work presented in this paper focuses on the development of an XML-based format, which extends the concepts of homeML [14], for storing data generated by a range of devices and services both within the home environment and also outside of the home.

A. XML

eXtensible Markup Language (XML) is a rule based language designed to transport and store data. It is one of the most common tools for data transmission and is becoming increasingly more popular in the area of storing and describing information [17]. The main benefits of using XML are its platform, vendor and application independence as well as offering an easy to follow hierarchical data structure [14]. XML is also considered to be simple to read and understandable by both users and computers.

There is an increasing desire to access and manipulate useful information from a multitude of data sources and achieve interoperability between a number of systems. Fengguang et al. [11] discuss
the use of XML as a means of storing data in a simple and rapid way, to enable the integration of data stored from various data sources.

Catley and Frize [18] discussed the potential application of a standardized medical database and integrated medical applications using XML along with the Unified Medical Language Systems (UMLS). Currently most medical databases are institution dependent and as such have a heterogeneous nature. It is hoped that through the use of XML, data sharing between institutions could be enhanced and challenges posed by non-existent medical database standardization could be solved.

IV. METHODS

The following section documents the evolution of homeML version 1.0 to homeML version 2.0. The validation process which was used is also presented.

A. Open-Source Data Sets

A number of open-source datasets are available online from various institutes that, have collected data generated from experiments conducted within intelligent environments [19].

<table>
<thead>
<tr>
<th>homeML</th>
<th>Date of Recording</th>
<th>Room Param</th>
<th>Room Description</th>
<th>Device Description</th>
<th>Device Location (X,Y,Z)</th>
<th>Device Type</th>
<th>Units</th>
<th>Real Time Information (runID)</th>
<th>Event (eventID)</th>
<th>Time Stamp</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>UUJ</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y (sensorID)</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y (status)</td>
</tr>
<tr>
<td>UA</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y (sensorID)</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y (val)</td>
</tr>
<tr>
<td>UTA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MavLab</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y (sensorID)</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y (status)</td>
</tr>
<tr>
<td>MavPad</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y (level)</td>
<td>Y (number)</td>
<td>Y</td>
<td>Y (info)</td>
<td>Y (zone)</td>
<td>N</td>
<td>Y (state)</td>
</tr>
<tr>
<td>WSU</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y (ID)</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y (value)</td>
</tr>
<tr>
<td>MIT</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y (sensor_ID)</td>
<td>N</td>
<td>Y</td>
<td>Y (zone)</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
</tbody>
</table>

Table 1. Results of homeML version 1.0 Review

The datasets used within this study included Massachusetts Institute of Technology (MIT) [20], University of Amsterdam (UA) [21], University of Texas at Arlington (UTA) [22], Washington State University (WSU) [22] and University of Ulster at Jordanstown (UUJ) [23]. These datasets were accessed in order to compare the data structure used by each to the homeML version 1.0, as shown in Table 1. The top row of Table 1 shows the various pieces of data that can be stored within the homeML format. Each dataset was compared to the format to identify any data that should was not included in the format but was required, as well as anything that could be considered
redundant. In addition, examination of the datasets were also useful in establishing data formatting requirements for revised versions of the schema.

homeML [14] was initially proposed to address the problems caused by the heterogeneous nature of data generated within smart home environments. Initially, homeML was critically reviewed to ensure the rationale and accuracy of the concept [14].

The review process, following consideration of the aforementioned datasets, identified a number of inaccuracies which required modifications to be made to the homeML schema. Table 2 provides an overview of the changes and rationale for each adjustment.

<table>
<thead>
<tr>
<th>Element</th>
<th>Alteration</th>
<th>Rational</th>
</tr>
</thead>
<tbody>
<tr>
<td>Room</td>
<td>‘RoomDescription’ now optional rather than essential.</td>
<td>Only 1 out of 5 research institutes recorded ‘RoomDescription’.</td>
</tr>
<tr>
<td>Room</td>
<td>‘Units’ now optional rather than essential.</td>
<td>No research institute recorded ‘Units’.</td>
</tr>
<tr>
<td>Ambulatory Device</td>
<td>This attribute is now optional rather than essential.</td>
<td>Although, a patient is living within a smart home environment they may not necessarily be required to wear a wearable device.</td>
</tr>
<tr>
<td>Ambulatory Device</td>
<td>‘Data’ was added to the ‘RealTime Information’ attribute.</td>
<td>Although, SampleRate, StartTime and EndTime were recorded, the data actually being produced by the device was not being stored anywhere.</td>
</tr>
</tbody>
</table>

Table. 2. Sample of minor alterations made to the original homeML schema.

B. homeML v 2.0

As technology and services continue to improve it is now possible to monitor a patient outside of their home environment and also deliver a limited number of assistive services. The data generated through the interactions with their environment will differ as the patient moves between locations. As a result it was necessary to further revise the homeML schema to support the storage of data generated within multiple environments to include both inside and outside of the home environment. The schema was extended further to incorporate the need to consider a patient's location when recording data. The hierarchical data trees can be viewed in Figure 1 overleaf. The data trees have also been
supplemented with Tables 3 and 4 which describe the elements and attributes within the aforementioned model.

As the hierarchical representation shows a device can be categorised as being either a location or a mobile device. Whereby, a location device is a device restricted to one location, whether it is integrated into the environment or not. An example of which would be a PIR movement sensor that is restricted to one location. A mobile device however, is not restricted to one location and can be easily moved between environments. An example of which would be an accelerometer, that is capable of being used continuously as a patient moves between environments.

<table>
<thead>
<tr>
<th>Location</th>
<th>The Location Element for XML based Smart Home Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Element/Attribute</td>
<td>Description</td>
</tr>
<tr>
<td>Location Description</td>
<td>General description of the location.</td>
</tr>
<tr>
<td>Location Class</td>
<td>Details the type of location, e.g. outside, smart home.</td>
</tr>
<tr>
<td>Location Device</td>
<td>Definition of location device.</td>
</tr>
</tbody>
</table>

Table 3. Description of HomeML element based on the hierarchical representation of homeML version 2.0 in Figure 1.

<table>
<thead>
<tr>
<th>homeML</th>
<th>The root element for XML based smart home data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Element/Attribute</td>
<td>Description</td>
</tr>
<tr>
<td>Inhabitant Details</td>
<td>Personal Information of Inhabitant.</td>
</tr>
<tr>
<td>Location</td>
<td>Definition of the Location and Devices within a Location.</td>
</tr>
</tbody>
</table>

Table 4. Description of Location element based on the hierarchical representation of homeML version 2.0 in Figure 1.
V. RESULTS AND DISCUSSION

It has been the aim of this study to develop a solution to the problems caused by the heterogeneous nature of data generated within a smart home environment. This study extends upon a proposed solution homeML [14] which aimed to develop a standard format for the storage and exchange of data. Upon close inspection the original homeML schema was considered to have a limited use in regard to the storage of data generated within multiple environments.

In our current work we have extended the concepts of homeML through assessing its exposure to a number of currently available online data sets in addition to assessing its ability to manage data gleaned from both inside and outside the home environment. In addition, a number of experiments were performed in order to generate new data that could also be input into the XML schema in order to validate whether the structure was compatible with a range of data sources and structures. The experiments involved tracking an individual within an environment using PIR and door sensors whilst wearing an accelerometer; as well as using a mobile phone (with built in GPS) to track an individual outside of the environment. Based on this validation the revised version of homeML proved to be adequate for all data requirements.

Further work will involve the development of an interface to allow the user to design the XML data schema to suit any dataset they may wish to record. The interface will eventually provide an end-to-end tool allowing users to both convert data to the homeML schema as well as providing a tool to visualise the data.
Figure 1. Hierarchical Representation of homeML version 2.0 as a series of Tree diagrams.

REFERENCES


Appendix I – NIBES’10 Abstract
An XML Based Format for the Storage of Data Generated within Smart Home Environments

H.A. McDonald, C. Nugent, G. Moore and D. Finlay

Computer Science Research Institute, University of Ulster, Northern Ireland

Correspondence to: swan-h2@email.ulster.ac.uk

Introduction

The proportion of the global population with an age greater than 60 is expected to increase from 10.0% in 2000 to 21.8% in 2050 and reach 32.2% by 2100 [1] [2]. Concurrently, the number of children being born is also predicted to decrease, with the ratio of 15 year olds to 65 year olds dropping from a current ratio of 9:1 to 4:1 by 2050 [3]. As a result, there will be fewer young people to assist older adults cope with the challenges of ageing [4].

As people grow older they are likely to experience age-related impairments; with the older population being the most frail and experiencing the highest incidence of disease and co-morbidity, as well as suffering a high level of disability [1].

All of these challenges caused by changing demographics will result in a major impact on the economy as the demand for health and social care services will continue to rise [5]. Although remaining at home can become challenging for an elderly person, many wish to remain in their own homes for as long as possible. Currently more than 80% of people 65+ are living independently [1].

Within the last decade efforts have been made to develop technology orientated solutions that can be used to monitor, diagnosis and manage patients’ conditions inside and outside of their home environment; particularly in the direction of moving a patient from a clinical venue, such as a hospital to their home. One of the many challenges within this domain is the issue of managing the heterogeneous nature of the data collected.

Materials and Methods

The data generated as a direct result of smart environments is largely heterogeneous. This is caused by it being recorded from a number of sensor based sources and services along with its storage in a wide variety of non-standardised formats. The main issue with the heterogeneity of the data is its lack of interoperability hence creating difficulties when data is being exchanged and also when it is being re-used and most importantly processed by a number of differing services.

As yet there does not exist a standard within the research community specifying how data generated within a smart home environment should be stored. Due to these issues there is a growing need along with an opportunity to develop a standard format for its storage and exchange.
The current work focuses on the development of an XML-based format by extending the concepts of homeML [6], for storing data generated by a range of devices and services both within the home environment and also outside of the home.

**Discussion**

It is the aim of this study to develop a solution to the problems caused by the heterogeneous nature of data generated within a smart home environment, extending upon the previously proposed solution of homeML which aimed to develop a standard format for the storage and exchange of data. Upon close inspection the original homeML schema was considered to have a limited use when storing data generated within multiple environments. The schema has been extended to resolve these limitations.

As a result of this work the homeML version 1.0 will be extended to incorporate the findings from our evaluation based on a number of sources of data

**References**


