

Domain Specific Modeling Language for Early Warning System: Using IDEF0 for Domain Analysis

Syed Imran, Franclin Foping, John Feehan and Ioannis M. Dokas

Cork Constraint Computation Centre, University College Cork
Western Gateway Building, Cork, Ireland

Abstract

Domain analysis plays a significant role in the design phase of domain specific modeling languages (DSML). The Integration Definition for Function Modeling (IDEF0) is proposed as a method for domain analysis. We implemented IDEF0 and managed to specify reusable domain specific language artifacts in a case study on water treatment plants safety. The observations suggest that IDEF0 not only examines the technical aspects of the domain but considers also its socio-technical aspects; a characteristic listed as DSML requirement in the case study.

Keywords: *Domain Analysis, Domain Specific Modeling, IDEF0, Reusable Domain Models, Early Warning Systems.*

1. Introduction

The objective of our research is to develop a graphical language able to represent concepts related to the domain of water treatment plants operations and safety. The graphical language will be an integrated component of an Early Warning System (EWS). The models specified by the graphical language will represent different facets of the domain and executable code will be generated automatically. In short, the objective of our research is to develop a Domain Specific Modeling Language (DSML) for EWS.

The first phase in the development of any DSML is to gather and process the required information and to define reusable language artifacts. This phase is known as domain analysis [1]. The idea of reusable artifacts is of importance. It forces the definition of filters based on which commonalities and variabilities of the domain become discernible.

During domain analysis, knowledge about the domain is acquired from various sources such as experts, technical documents, procedures, regulations and other materials. The purpose of this phase is to produce domain models. Domain models consist of concepts, relations, terminology and semantics representing the overall knowledge of the domain and hence revealing important properties that affect the design of the DSML [2].

The problem is that there are not any lucid guidelines or standards on how to gather and process the information of the domain when developing a DSML [2,3]. Formal analysis techniques such as the Feature-Oriented Domain Analysis (FODA) [4], Family-Oriented Abstractions, Specification and Translation (FAST) [5], Domain Analysis and Reuse Environment (DARE) [6] have been proposed in domain engineering. These approaches aren't so popular among developers of DSML. An explanation for this is that domain engineering is strong on its main focus -finding and extracting domain terminology, architecture and components- but gives little help in designing, constructing and introducing languages for the engineered domain [7].

In this paper we propose the use of Integration Definition for Function Modeling (IDEF0) as part of a domain analysis method for a DSML. The DSML is developed to become integrated component of an EWS. A case study taken from water treatment plant operations and safety domain demonstrates its usefulness.

The remainder of the paper is structured as follows: In Section 2, the IDEF0 method is briefly described followed by a description of the steps taken to curry out the domain analysis in the case study. In Section 3, we demonstrate first the implementation of IDEF0 in defining the entities and processes of the Water Treatment Plant (WTP) and subsequently in identifying the social factors that affect the operations of WTP. In Section 4, the reusable domain model is briefly described followed by the conclusions.

2. Methods

2.1 IDEF0

IDEF0 is designed to model the activities, processes and actions of an organization or system [8]. It is an engineering technique for performing and managing needs analysis, benefits analysis, requirements definition, functional analysis, systems design, maintenance, and baselines for continuous improvement [9]. It enhances the domain expert involvement by through simplified graphical representation.

With IDEF0 one can specify five elements: the activities, input, output, constraints or control and mechanisms or resources of a system (see Fig. 1). The activities are receiving certain inputs that need to be processed by means of some mechanism, and are subject to certain control such as guidelines, policies, before being transformed into output. The activities can be instantaneous or can happen over a period of time.

IDEF0 diagrams are organized in hierarchical structure, which allows the problem domain to be easily refined into greater details until the model is descriptive as necessary. The first level of the hierarchical structure represents a single high level activity, which then can be decomposed to lower levels that represent the processes of the system, the resources and the information that passed between them in more detail.

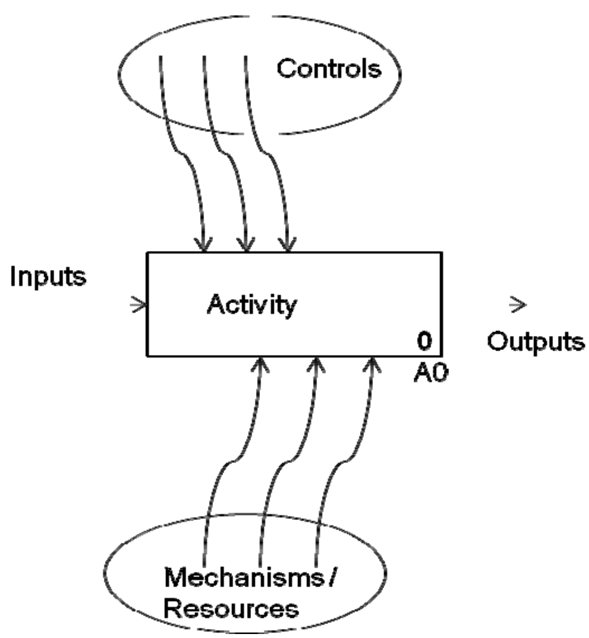

Fig. 1 IDEF0 Language model.

Some of the characteristics of IDEF0 are [9]:

- It provides comprehensiveness and expressiveness in graphically representing a wide variety of business, manufacturing and other types of enterprise operations to any level of detail;
- It provide concise details in order to facilitate its understanding, consensus and validation;
- It is a coherent and simple language and promotes consistency of usage and interpretation;
- It enhances communication between system analysts, developers and users through ease of learning and its emphasis on hierarchical exposition of detail.

IDEF0 has been extensively used for structured analysis and design of systems, covering functional modeling in describing the system graphically by means of providing better understanding. It has also been used for many reasons such as for designing and specification of methodologies in which progressively decomposing a system from a high level was found to effective for methodology development [10]. It is widely used also in modeling manufacturing systems in which knowledge based system approach is used [11].

2.2 Steps of Domain Analysis and Case Study

During the domain analysis phase we visited the Inniscarra WTP in County Cork, Ireland on several occasions for over a three month period. The Inniscarra plant serves the population of about 100,000 consumers and is distributing about 21,000 m3 /day of treated drinking water. It is considered as one of the largest WTP in Ireland. With the help and guidance of the domain experts it was possible to gain enough information for the analysis and to build IDEF0 models.

The requirements of the overall modeling were agreed over a series of meetings with domain experts. Once gathered, the requirements were documented and validated to insure that there is no inconsistency or obscureness. Some of the key requirements are briefly described below.

- Systematic profiling should be included as a feature of the DSML. With the term “systematic profiling” we mean the modeling and representation of physical entities that constitute a WTP in a categorized way. For example, many physical objects are involved in each purification phase of raw water such as flocculation, coagulation. The goal of the systematic profiling is to provide detail information of these objects, categorized by purification phase in which they belong. The profiling process guides us to understand what exists in a WTP providing basic information of its capacity and capability to perform specific purification tasks. For example, if a WTP has an ozone generator in its profile we can assume that the plant design to perform "Ozonation" process to disinfect the water, to reduce color and odor by oxidizing the organic substances in the water.

- The DSML has to have a hazard/vulnerability component permitting its users to represent their mental models about the threats and vulnerabilities of the WTP, providing thus an explanation on how unwanted events may happen during WTP operations.

- Users should be able to define early warning signals. There are two categories of warning signals; namely weak and sound warning signals. For example, sensors monitoring the water for the presence of chemical radiological or biological hazards are able to provide sound warning signals. Weak signals are those which indicate the presence of latent conditions in a WTP. In short, are indications of how much a WTP is vulnerable to the variability of the environment. For example, if a WTP is constructed close to the banks of a river, then its location may be considered as a latent condition which makes it vulnerable to river floods incidents. Usually, latent conditions are detected by the personnel or by the auditors of the WTP.

- Users should be able to define socio-technical factors that may affect the safety of WTP. By socio-technical factors we mean the social and organizational features that may affect WTP operations.

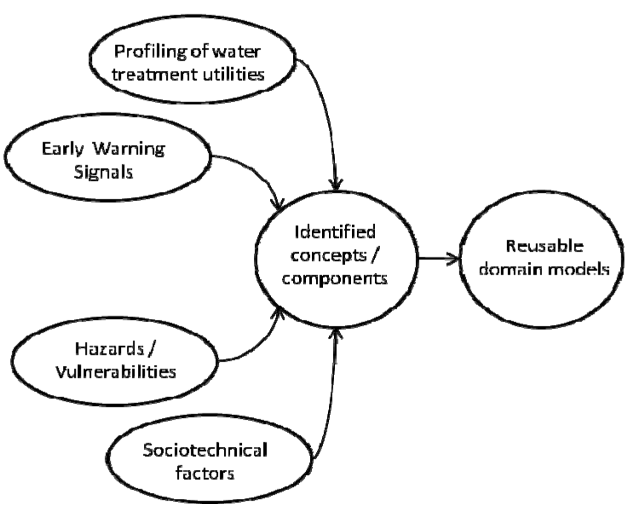

Fig. 2 Reusable domain models for EWS in WTP.

Fig. 2 illustrates an overall view for developing the reusable domain models. A profiling process performed first to list the physical entities and processes of WTP. In this step we made extensive use of IDEF0. It is the step where this paper is focusing at. In a second step, hazards, vulnerabilities and early warning signals together with the social factors that affect the safety health of WTP were defined based on system safety engineering approaches. Unexpectedly, as it is shown in Subsection 3.1, during the first step we found out that IDEF0 helped us in identifying important social factors that effected WTP operations. We then mapped the concepts derived from both, the first and the second steps and defined their relations and constraints. Finally we specified the reusable artifacts that constitute the DSML of the EWS.

3. IDEF0 model of WTP

The high level view of the WTP is illustrated in Fig. 3. The model describes the entire WTP as an activity in which raw water is taken as an input. The activity is executed using different resources such as chemicals, treatment mechanisms and labour. Controls ensure that all WTP activities accord with legislations and standards. The output of the activity is the treated water.

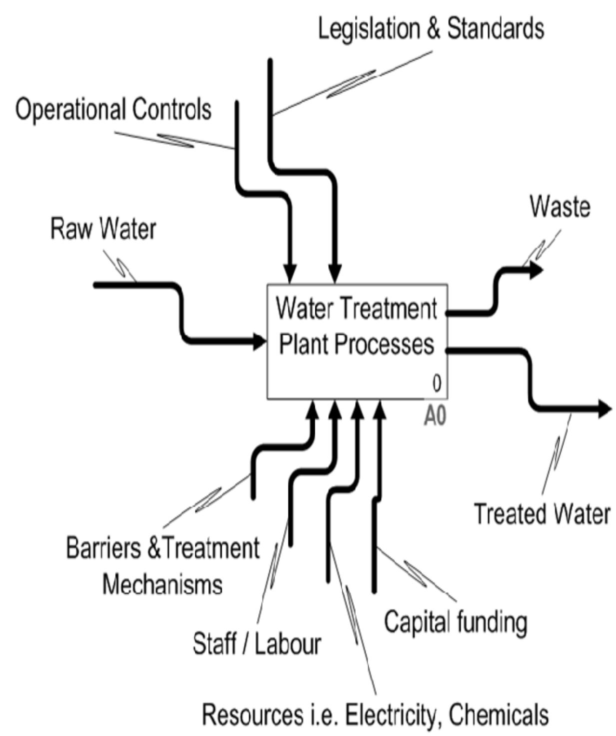

Fig. 3 High level IDEF0 view of the WTP.

Water purification consists of several stages such as water intake and screening, flocculation, coagulation, filtration, pre and post chemical treatment and disinfection. The IDEF0 model represents these steps through hierarchical decomposition. A part of the lower level view is shown in Fig. 4 where the higher level activity of the "WTP process" is broken down into several lower level activities resulting to better understanding of water purification processes.

The first activity in Fig. 4 is the "Intake & Storage Process". Raw water intake is the input of this activity. The raw water is then handled by mechanisms such as penstock, intake pumps, metal screens. Furthermore, the activity is performed under some controls such as setting the penstock selection level, regulating the intake pumps, checking of raw water quality parameters before allowing the raw water to be passed into next process called "Pre-Filtration Process". The sequence of the water flow is from left to right and each activity is decomposed further to several lower levels so as to provide the obscure understanding of the WTP treatment processes.

As result, with the IDEF0 analysis of the WTP we identified 21 activities in 4 decomposition levels. In the last decomposition level we identified 27 mechanisms and 25 controls. This indicates the usefulness of IDEF0 in:

- Gaining comprehensive information on the process involved in WTP;

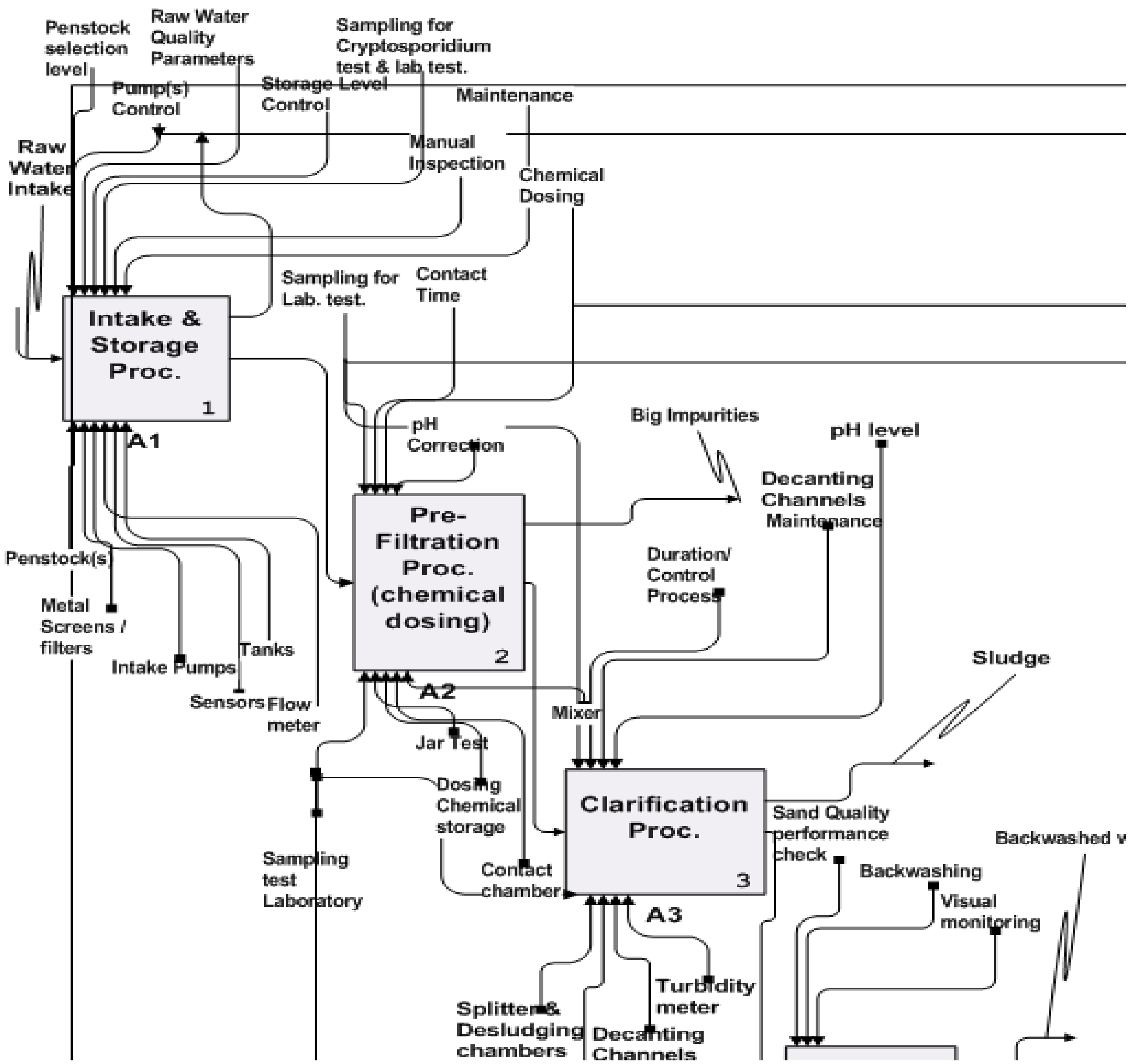

Fig. 4 IDEF0 model for A0 activity in WTP treatment

- Identifying the output flows between activities which are integral to understanding critical elements for safety and security;
- Identifying critical concepts;
- Listing the physical objects of WTP.

3.1 Socio-technical factors

Several social factors such as the availability of funds, the regulations of the governance body on different standards, the strategic actions taken in case of flooding or other natural disasters, may affect the water treatment process. Table 1 illustrates how IDEF0 facilitate in understanding these factors. It lists some of the socio-technical factors associated with the mechanism “Intake pumps” of activity A0-1 (see Fig. 4). The table provides the information regarding the technical and socio-technical factors that can affect the operation of the water pumps in supplying water to the specific process of the WTP. It also provides us with the information on how socio-technical factors relate to some of the technical factors that could lead in affecting the entire water intake and storage process. For example, when the demand of the drinking water is higher than the original water treatment capacity of the plant, the social factors plays the major role in deciding how to manage the situation rather to allocate additional resource of water supply or the funds to expand or to upgrade the whole or part of the water treatment facility or in case the funds are not accessible the additional burden is put on the treatment process hence effecting the operation of water intake pumps, storage tank level, raw water quality parameters.

4. Reusable Domain Model

The reusable domain model is specified using a metamodel. The metamodel is often defined as a model of a model that is the language in which the models are defined and they themselves also needs to be expressed in some language, for this purpose we have used model-based Eclipse Modeling Framework (EMF) technology [12]. The EMF is originally based upon a subset of the Object Management Group standard Meta Object Facility 2.0 (MOF) called Essential MOF (EMOF) [13]. The development of metamodel has been assisted by the information obtain from the IDEF0 model. During the development phase, as a part of prerequisite the information obtained from the IDEF0 is categorized according to the commonalties and variabilities, where commonalties define the set of features that are common to several WTP activities and variabilities are what makes activities different from each other.

Table 1: Socio-technical factors from IDEF0 Analysis

IDEF0 model highlighting technical and socio-technical factors in regulating raw water intake pumps.

Activity: A0-1
Name: Intake and Storage process
Mechanism: Intake pumps
Control: Pump control

Description: The raw water intake pumps takes water from a lake and supplies it to the raw water reservoir in WTP. This activity involves control measures, which are influenced by different technical and subsequently socio-technical factors.

Technical:

1. Feedback from activity A5 (Treated water outlet), information of the treated water level from treated drinking water reservoirs.
2. Raw water quality parameters.
3. Raw water intake level.
4. Raw water storage tank level.
5. Maintenance and other technical reasons.

Socio-technical:

- The treated water supply demand should be kept in consideration to the WTP capacity. If the demand of drinking water is more than the treatment capacity of WTP then the technical factor “1” will be affected as the level of drinking water reservoir will mostly be low and would ultimately effect the pump control action in regulating raw water intake pumps.
- Raw water quality parameters are set by national and EU standards and legislation.
- Raw water intake level is determined not only by the water resource but also depends on the level of water in the lake. As in our case study, the lake water level up to which limit the water can be taken into WTP is set by Water Services Authority. There is hydroelectric plant downstream from our water intake, which requires subsequent amount of water to be fully functional.
- Raw water storage tank level depends on many factors such as, the availability of water and the demand of treated water in supply. In some cases raw water storage tanks are not available because of financial limitations.
- Funds allocated to run water treatment utilities may have significant effects on their performance. Maintenance, periodical inspections and audits are required.

Table 1: Commonalities and Variabilities using IDEF0 model

Activity	Sequence No	Subsequence No.	Mechanisms		Controls	
			Commonalities	Variabilities	Commonalities	Variabilities
Intake	A0-1	A1	1. Flow meter 2. Tanks 3. Sampling test laboratory	1. Penstock 2. Metal Screens 3. Intake pumps	1. Pumps Control 2. Manual Inspection 3. Maintenance 4. Sampling for cryptosporidium test and laboratory test.	1. Penstock selection level. 2. Raw water quality parameters

The mapping from IDEF0 to metamodel is illustrated in Fig. 5. During the mapping process, all the activities are defined as separate class each, which also act as superclass to their subclasses defined for their sub activities. For example, the first activity of our IDEF0 model is “Intake and Storage Proc” which has been mapped to class “Intake” and it is also act as a superclass to the sub-activities specified at its in its lower decomposition level that are “Raw Water Intake Proc.” and “Raw Water Storage Proc.”.

The classes are also specified for “mechanisms” and

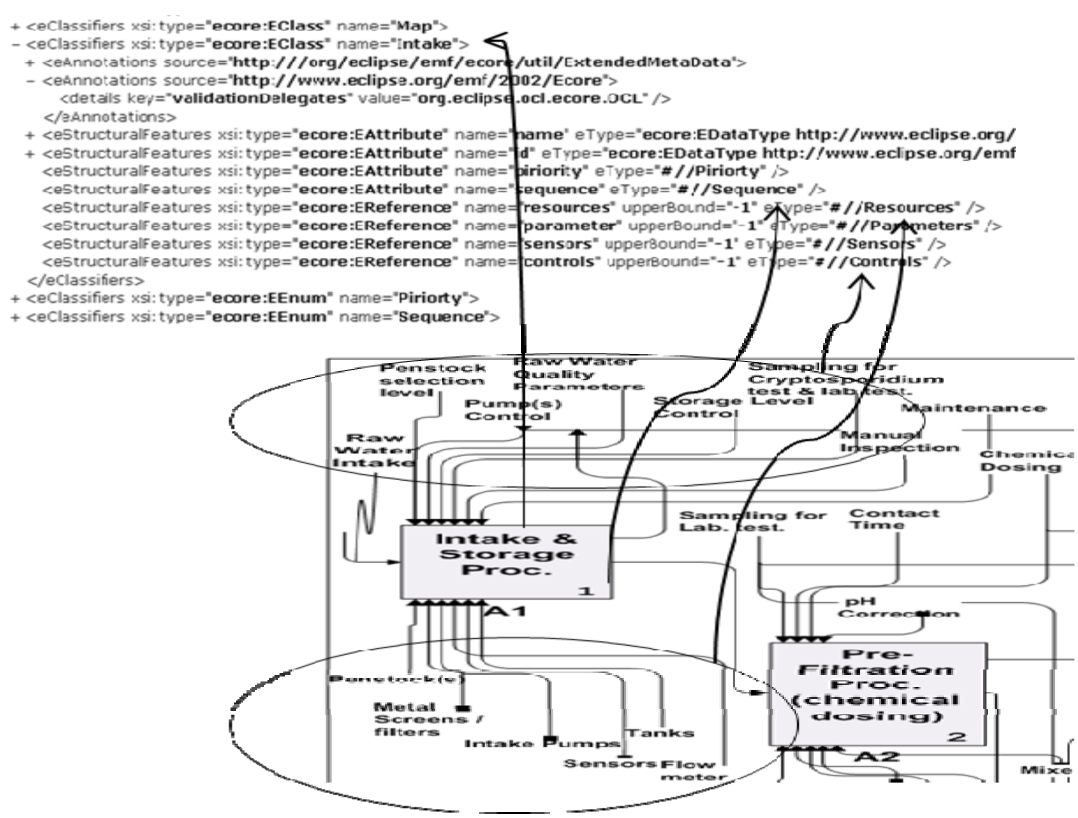

Fig. 5 IDEF0 model for A0 activity in WTP treatment process.

"controls", each having subsequent subclasses representing their "variabilities". Moreover, "commonalities" are defined in a separate class. It is also to be noticed that commonality can be define as an attribute of variabilities but the vice versa is not allowed. For example, the class "Penstock" which is categorized as a variability of a mechanism in an "Intake" activity can have an attribute class "Sensor" but the other way around is not possible.

Once the required classes are defined they are referenced to each other accordingly as defined by the IDEF0 model. A class with subsequence number A1 is to be a referred by class with a sequence number A0-1 as it is the sub activity of A1. The "sensor" class is referred by the "mechanism" class which is further referenced by the "category" class. Here the sensor is defined as commonality because it is repeated in many sub activities that belong into a lower hierarchy level of the "Intake and Process" shown in Fig 4. The activities in the lower hierarchy levels made use of different types of sensors. Given this repetition we defined the "sensor" concept as commonality. Fig 4 shows the 2nd level of the IDEF0 analysis to demonstrate the result of IDEF0 method. In reality we developed a series of hierarchical IDEF0 models that describe each activity box shown in Fig 4 in more detail. All classes are referred by the "Map" class in our case. In EMF's ecore metamodel the class "Map" is considered as a starting point or mapping point to rest of the metamodel. The class "Map" itself does not take part directly in modeling of domain concepts but it is used as a point where the whole architecture of the metamodel is initially mapped as referenced. In the case of "controls", we used a declarative language in a metamodel level, the object constraint language, and also specify separate "control" class, to define their constraints.

5. Conclusions

This paper presents the use of IDEF0 as domain analysis approach for the development of a DSML. The case study demonstrated that IDEF0 contributed in identifying the reusable building blocks of a DSML and proved very useful in collecting, organizing and representing the relevant domain information.

On the other hand, it was observed that IDEF0 models can become so concise that it can cause comprehension difficulties and only domain experts can understand it fully. Difficulties can also arise in communicating various domain concerns between different domain experts. Furthermore, it has been noticed that the domain information can change during the course of time such as due to advancement of technology or other issues; in this case there will always be a need to update and validate such changes in IDEF0 model. We clearly understand the limitations of using IDEF0 model for socio-technical aspects as it does not allow us to model the interactions between social factors. However, it guides us to identify parameters which are more influenced by it.

There is no standard methodology for domain analysis that fits all domains due to conflicting goals and objectives. Therefore, there is always a need to understand which approach for domain analysis fits the needs and best matches with the specifications. This paper demonstrated that IDEF0 is advantageous in defining the physical objects and the processes in a complex system but also contributed in identifying other social factors that affect the performance of the system. These characteristics are of importance for any EWS. In short, IDEF0 should be considered as a reliable alternative approach for the development of a DSML in EWS.

Acknowledgments

This work is supported by the research project SCEWA (Grant No 2007-DRP-2-S5), funded by the Irish Environmental Protection Agency under the DERP grand scheme. We would also like to express special thanks to Mr. Pat Murphy, the plant managers from Inniscarra Water Treatment Plant for his cooperation and assistance.

References

[1] J. M. Neighbors, "Software construction using components", doctoral dissertation, Univ of California, Irvine, 1981.
[2] R. Tairas, M. Mernik, and J. Gray, "Using ontologies in the Domain Analysis of Domain Specific languages" Models in Software Engineering. Vol. 5421, 2009, pp. 332-342.
[3] M. Mernik, and J. Heering, and A.M. Sloane, "When and how to develop domain-specific languages", ACM Computing Surveys (CSUR), Vol. 37, 2005, pp. 316-344.
[4] K. Kang, S. Cohen, J. Hess, W. Nowak, S. Peterson, " Feature-Oriented Domain Analysis (FODA) Feasibility Study", Technical Report CMU/SEI-90-TR-21, Software Engineering Institute, Carnegie Mellon University, Pittsburgh, PA, 1990.
[5] D. M. Weiss, C. T. R. Lai, "Software Product-Line Engineering: A Family Based Software Development Process", Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA, 1999.
[6] W. Frakes, R. Prieto-Diaz, C. Fox, "DARE: Domain Analysis and Reuse Environment", Annals of Software Engineering, Vol. 5, 1998, pp. 125-151.
[7] J.-P. Tolvanen, "Keeping it in the family", Application Development Advisor, 2002.

[8] P.S. DeWitte, R.J. Mayer, and M.K. Painter, "IDEF family of methods for concurrent engineering and business re-engineering applications", Knowledge Based Systems. College Station, TX, USA, 1992.

[9] FIPS, "Integeration Definatin of Functional Modeling (IDEF0)", Federal Information Processing Standards. Publication 183, Springfield, VA, USA, 1992.

[10] A. Presley, and D. Liles, "The Use of IDEF0 for the Design and Specification of Methodologies", Proceedings of the 4th Industrial Engineering Research Conference, Nashville, USA, 1995.

[11] C. L. Ang, R. K. Gay, L. P. Khoo, and M. Luo, "A knowledge-based Approach to the Generation of IDEF0 Models", International Journal of Production Research, Vol. 35, 1997, pp. 1385-1412.

[12] Eclipse Modeling Framework (EMF), http://www.eclipse.org/emf

[13] Essential MOF (EMOF) as part of the OMG MOF 2.0 specification, http://www.omg.org/docs/formal/06-01-01.pdf.

Syed Imran is a postgraduate student at University College Cork, Ireland. His research is focused on the development of a Domain Specific Modeling language for Early Warning Systems. He has masters in computer science from Technical University Braunschweig, Germany and bachelors in engineering from Voronezh State University, Russia. He has a working experience of more than 7 years as a software engineer.

Dr. Ioannis M. Dokas has diploma in civil engineering and PhD in expert systems. He has been awarded with EU Marie Currie Fellowships two times, during which he conducted post doctoral research in Germany and Ireland. Currently he is the leader of a research team working on early warning systems at the Cork Constraint Computation Centre. His current research is funded by the Irish EPA. His research interests are in the fields of systems safety, environmental engineering, and artificial intelligence.

John Feehan is a graduate of the Sligo Institute of Technology, Sligo, Ireland where he completed an Honours Degree in Environmental Science and Technology in 1990. Upon graduation he worked as an Environmental Manager in the Pharmaceutical Industry for five years before joining the newly established Environmental Protection Agency (EPA) in 1995. In the EPA, he has worked on strategic management, policy development, licensing, enforcement, surface water, wastewater and drinking water. John is also a graduate of University College Cork, Ireland where he was awarded a first class honours Masters Degree in Environmental Analytical Chemistry in 1996. He is Member of the Royal Society of Chemistry, the Chartered Institution of Water and Environmental Management and the Institution of Engineers of Ireland. He is a Chartered Chemist, Chartered Water & Environmental Manager, Chartered Scientist and a Chartered Environmentalist.

Franclin Foping is currently undertaking a Ph.D. in Computer Science at the University College Cork in the Republic of Ireland. The bulk of his thesis is to deploy an early warning service over the Internet. He also holds an MSc in Computer Graphics Programming from the University of Hull in England. Franclin was a finalist of the Microsoft Imagine Cup Software Design in March 2007 in England. He is a Linux and open source enthusiast with an experience of over 10 years in the field. In his spare time, Franclin enjoys travelling, reading, watching soccer and tennis games.