# Taikichiro Mori Memorial Research Grants

# Graduate Student Research Development Grant 2016

# Report

**Research project:** A Multi-dimensional Semantic Space Creation of Heavy metals in Water Quality Analysis System for Interpreting Environmental Situations and Real-time Notification by rSPA Processes.

**Researcher:** Chalisa Veesommai

**Affiliation:** Graduate School of Media and Governance

**Type of Program:** Doctoral Program (2nd years)

**Email Address:** chali-h@hotmail.com, chalisav@sfc.keio.ac.jp

**Research Advisor**: **Yasushi Kiyoki**

**Email Address:** kiyoki@sfc.keio.ac.jp

**Results of this project Research in Published Journal:**

Veesommai C., Kiyoki Y., Sillberg P., Soini J., Jaakkol H. and Chawakitchareon P. 2016. "The rSPA Process Realization: The Creation of River Heavy Metal Evaluation Index (rHMEI) by Using Dimensional Subspace of Heavy Metal". 2016 International Transaction Journal of Engineering, Management, & Applied Science & Technologies, Vol 7 (3), 189 – 203. (ISSN 2228-9860, eISSN 1906-9642).

**Future Challenges:**

- Symbolic-filtering and semantic-ordering functions for the analysis of water quality in wide areas
- The difference-extraction computing for water-quality analysis

International Transaction Journal of Engineering, Management, & Applied Sciences & Technologies

http://www.TuEngr.com

# The rSPA Process realization: The creation of river Heavy Metal Evaluation Index (rHMEI) by using dimensional subspace of heavy metal

Chalisa VEESOMMAI [a*], Yasushi KIYOKI[a], Pekka SILLBERG[b], Jari SOINI[b], Hannu JAAKKOLA[b] and Petchporn CHAWAKITCHAREON[c]

[a] *Graduate School of Media and Governance, Keio University, JAPAN*
[b] *Department of Information Technology, Tampere University of Technology, FINLAND*
[c] *Department of Environmental Engineering, Faculty of Engineering, Chulalongkorn University, THAILAND*

**ARTICLEINFO**

*Article history*:
Received
Received in revised form

Accepted
Available online

*Keywords*:
River Water Quality Analysis,
Database,
Multi-dimensional subspace of heavy metal,
river Sensing Processing Actuation (rSPA) processes,
river Heavy Metal Evaluation Index (rHMEI)

**ABSTRACT**

Several implementations in water-quality field are published and the important direction of the result doesn't give the overall of water quality in parts of heavy metal and lead to uncomplicated for public utilization as a specific word. It's necessary to realize the tool and processes for analysis system by creating the evaluation index and applying multi-dimensional subspace for minimized limitation. The river Heavy Metal Evaluation Index (rHMEI) on river Sensing Processing Actuation processes is created by using the multi-dimensional space of heavy metal substances, and applied to Pori's water resource (Finland) and evaluated an effect of nine heavy metal parameters. The rHMEI is feasible and effective for analyzing water quality in several categories. In the implementation of the analysis system, we integrate special knowledge resources in environmental analysis and semantic computing for evaluating water quality in heavy metal parts and interpreting numerical values of heavy metal to feature semantic wording.



## 1 Introduction

The water resource is significant in human life to be used for several kinds of consumptions. Generally, the different kinds of pollutants as biological chemical and physical characteristic are contained in water resources, which cause heavy effects on water born diseases and

*Corresponding author (C. Veesommai). Tel/Fax: +81 80 8863 9389. E-mail addresses: chalisav@sfc.keio.ac.jp, chali-h@hotmail.com  eISSN:
Online Available at

several problems on water resources, such as water shortages, pollutant contamination and accumulation in water resources and wastewater (Veesommai C. *et al.*, 2016). One of the main considerations on a water quality area is the heavy metal substance as chemical characteristic, which causes toxic effects even at low concentration (Das A.K., 1990). The pollution of water resources with heavy metal has caused a major threat to the human life to series hazard, disease by food chain and expose. These heavy metals make toxicity effects for a living organism (Mona H.M., 2014), persistent environmental contaminants and accumulate into the tissue or organism of aquatic vegetation and aquatic life. The heavy metals are into the natural cycle of water resources by exploitation of natural resources, as energy exploration, mining and basic human activities in agriculture, industrialization, and urbanization (Veesommai C. and Kiyoki Y., 2015).

In the field of environment engineering, it is pointed that one of the main cause of the serious problem is water contamination. Many organisms in the ecosystem are damaged by unclean water and inadequate sanitation. In this paper, we are focusing on several types of heavy metal contamination in water resource and realizing the analysis method by using semantic computing and subspace selection in part of the heavy metal parameter.

Semantic computing and subspace selection

Currently, the semantic computing is effective in several fields of research. Semantic space is useful for analyzing the results of multiple attributes on semantic-based analysis as database and data mining, analysis on the social marketing, and the environmental field. Several significant research results have been created in the semantic computing to realize semantic spaces or/and multi-spaces to analyze the meaning of words, sentences, numeric values and simulation of environment change (Kiyoki Y. *et al.*, 2015), and also the semantic computing has been applied to biological, chemical, GIS system as 5D World Map System, medicine and music field (Kiyoki Y. and Chen X., 2015), created the new method and the system of meta-database for extracting appropriate images referring to impress and content of image by using the Mathematical Model of Meaning (MMM) (Kiyoki Y. *et al.*, 1994) , which has proposed a metadata-base knowledge system with a new search method of semantic associative search based on the MMM. The MMM has been applied to the medical semantic spaces in the medical knowledge filed with semantic space integration and a domain-specific semantic associative searching (Kiyoki Y. and Kawamoto M., 2007).

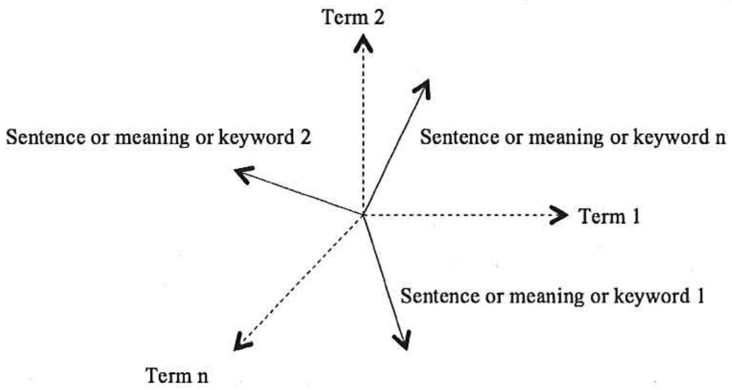


Figure 1: The semantic space.

Figure 1 shows the overview of the semantic space. The results of multiple attributes are created on semantic-based analysis in this space with the words, sentences or numeric values. The concept of the semantic space is promising for multiple attributes and we apply it to the analysis for multi-heavy metal in water resources on processing part of rSPA processes.

rSPA processes

The river Sensing Processing Actuation (rSPA) process was realized from the SPA concept (Kiyoki Y. and Chen X., 2015). The river Sensing Processes Actuation is a supporting concept for the definition of rules for automatically performing it, which is defined by Veesommai C. and Kiyoki Y. The meaning of sensing is the part of input data of water-quality as multi-parameters. The meaning of processing is some parts of analysis and knowledge databases. The meaning of actuation is the part of output action for notification and warning for sending action and updating the information of current situation. On the other hand, Several implementation and tool are creating or developing for analysis of water quality as an environmental index. The limitation of results from implementation is complicated understanding for public utilization, the limit in specifying target user and parameter, the limit in accuracy as an eclipsing and ambiguity of the environmental index result.

The problems of this study are: (1) how to analysis and summarize water quality in several parameters into one evaluate index, (2) how to create the multi-dimensional subspace of heavy metal and integrate the special knowledge resources in environmental analysis and semantic computing for evaluating water quality.

*Corresponding author (C. Veesommai). Tel/Fax: +81 80 8863 9389. E-mail addresses: chalisav@sfc.keio.ac.jp, chali-h@hotmail.com  eISSN:
Online Available at

In this paper, we propose a method for analyzing the quality of water resources, depending on the type of data samples, type of target groups, the size of the samples and informational goals. The index is one of the effective ways to present the results of the estimation-related environmental situation by several parameters or attributes. In addition, we also create the index from the dimensional subspace of heavy metal for estimating the quality of water resources in term of heavy metal in processing some parts on rSPA processes.

The essence of this paper is to illustrate the meta-level knowledge of the database system for the environmental engineering field in water resources and rivers. We propose a method for classification and interpretation of monitoring data with river Heavy Metal Evaluation Index (rHMEI) by using the dimensional subspace of heavy metal on rSPA processes.

The context in section 2, we collected the historical data of water quality in Pori' water resource, Finland shown as subsection 2.1 and 2.2. The methodology for realizing tool and the process of analysis system shown as subsection 2.3. The analysis part of rSPA processes has 2 procedures; the first procedure of processes performs the processing part (P) of rSPA processes with index creation (rHMEI) by using multi-heavy metal parameters show as subsection 2.3.1. Second procedure of processes analyzes data by using an index on multi-dimensional subspace shown as subsection 2.3.2. In section 3, we present the result of the index creation (rHMEI) processes and implementation shown as subsection 3.1 and the result of data analysis shown as subsection 3.2. In section 4, we summarize the result of this paper.

# 2 Material and Method

## 2.1 The study area and description and Data Structure

The collected water quality data in this study is utilizing Open Data (Open Knowledge Foundation, 2016) of surface water quality provided by Finnish Environment Institute SYKE (SYKE, 2016). In 2015, the dataset contained over 2.7 million water samples and 28 million analysis results from almost 70000 locations all over Finland. For this study, we focused on sampling points located in the municipality of Pori. The population of Pori is 85000 people. The city of Pori is located on the west coast of Finland on the estuary of Kokemäenjoki river about 15 kilometers from the Gulf of Bothnia.

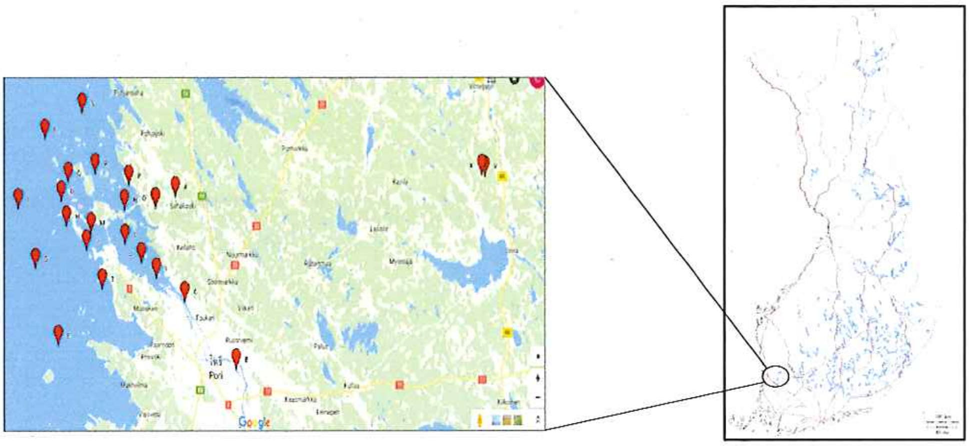

Figure 2: The geographic location and sampling points of data.

The sampling locations included in our study are shown on the map in Figure 2, and corresponding names and exact coordinates of them are listed in Table 1. The locations were selected as they were updated on a regular basis (at least 4 samples peryear), exact coordinates (latitude, longitude) from 61.64659, 21.68996 to 61.66722, 22.50495 in the time span of 2013 - 2016. The number of sampling locations is 23 points as follow: one from Eteläjoki (river), one from Metsä-Ahla (water basin), two from Kokemäenjoki (river; abbreviated Kojo), two from Rahkakeidas, and 17 from the Gulf of Bothnia in front of Pori (abbreviated Pome).

Table 1: The description of the area of rivers and the sets of points.

| Point | Name of points | Location | |
|---|---|---|---|
| | | Latitude | Longitude |
| A | Eteläjoki tie 272 | 61.64659 | 21.68996 |
| B | Kojo 37 Varvoorinjuopa | 61.48256 | 21.85145 |
| C | Kojo 46 Isojuopa | 61.54688 | 21.71592 |
| D | Metsä-Ahla allas MA1 | 61.63602 | 21.63688 |
| E | Pome 119 Iso-Ensk luot | 61.70105 | 21.34225 |
| F | Pome 235 Säppi koill | 61.50524 | 21.37808 |
| G | Pome 260 Mkallo 4 mpk lo | 61.57857 | 21.31642 |
| H | Pome 270 Reposaari lä | 61.61858 | 21.39974 |
| I | Pome 276 Hylkiriutta lo | 61.63525 | 21.27138 |
| J | Pome 50 Pussaanluoto | 61.56925 | 21.63852 |
| K | Pome 51 Sådösaar et | 61.58369 | 21.60059 |
| L | Pome 56 Kolppa | 61.60155 | 21.55666 |
| M | Pome 58 eteläselkä | 61.61230 | 21.46563 |
| N | Pome 64 Lannask koill | 61.63422 | 21.55499 |
| O | Pome 67 Tahkol luot | 61.64261 | 21.38593 |
| P | Pome 70 Kristisk lä | 61.65742 | 21.56636 |
| Q | Pome 71 Arvenk pohj | 61.66005 | 21.40482 |
| R | Pome 72 Iso-Väkk lä | 61.66861 | 21.47688 |
| S | Pome 83 Isot Plokit lä | 61.72599 | 21.44163 |
| T | Pome 86 Yyterin ed | 61.55881 | 21.49494 |
| U | Pome 88 Kolmikulma | 61.59592 | 21.45334 |
| V | Rahkakeitaan kp kaiv 10 | 61.66603 | 22.51338 |
| X | Rahkakeitaan kp oja itä | 61.66722 | 22.50495 |

*Corresponding author (C. Veesommai). Tel/Fax: +81 80 8863 9389. E-mail addresses: chalisav@sfc.keio.ac.jp, chali-h@hotmail.com 

## 2.2 Data structure

We have collected several metal and heavy metal data, such as Aluminum (Al), Arsenic (As), Barium (Ba), Cadmium (Cd), Chromium (Cr), Copper (Cu), Mercury (Hg), Lead (Pb), Nickel (Ni), Iron (Fe), Titanium (Ti), Vanadium (V) and Zinc (Zn) from 23 points with added spatiotemporal metadata, such as location and date for each data in the data structure shown in Figure 3.

| id | Al | Ba | Ar | Cd | Cr | Cu | Fe | Pb | Ni | Zn | hg | Ti | V | ndate | Location | Latitude | longitude |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 450 | 0 | 0.72 | 0.04 | 1.2 | 2.4 | 1800 | 0.45 | 4.5 | 7.9 | 0 | 0 | 0 | 1/7/13 | 66209::EtelŠjoki tie 272 mts | 61.64659 | 21.68996 |
| 2 | 340 | 0 | 0.75 | 0.02 | 1.1 | 2.3 | 1900 | 0.4 | 4.1 | 5.6 | 0 | 0 | 0 | 2/12/13 | 66209::EtelŠjoki tie 272 mts | 61.64659 | 21.68996 |
| 3 | 380 | 0 | 0.71 | 0.02 | 0.8 | 2.1 | 2600 | 0.48 | 3.4 | 10 | 0 | 0 | 0 | 3/12/13 | 66209::EtelŠjoki tie 272 mts | 61.64659 | 21.68996 |
| 4 | 680 | 0 | 0.73 | 0.05 | 0.87 | 2 | 2100 | 0.9 | 4.9 | 6.5 | 0 | 0 | 0 | 4/9/13 | 66209::EtelŠjoki tie 272 mts | 61.64659 | 21.68996 |
| 5 | 280 | 0 | 1.2 | 0.06 | 1.5 | 2.8 | 4500 | 1.5 | 3.5 | 15 | 0 | 0 | 0 | 4/18/13 | 66209::EtelŠjoki tie 272 mts | 61.64659 | 21.68996 |
| 6 | 1800 | 0 | 1 | 0.047 | 3.7 | 3.5 | 2700 | 1 | 4.4 | 11 | 0 | 0 | 0 | 4/25/13 | 66209::EtelŠjoki tie 272 mts | 61.64659 | 21.68996 |
| 7 | 680 | 0 | 0.82 | 0.034 | 0.92 | 1.8 | 1600 | 0.72. | 3.5 | 12 | 0 | 0 | 0 | 5/6/13 | 66209::EtelŠjoki tie 272 mts | 61.64659 | 21.68996 |
| 8 | 280 | 0 | 0.74 | 0.031 | 1.2 | 2.8 | 1600 | 0.66 | 3.6 | 5 | 0 | 0 | 0 | 5/21/13 | 66209::EtelŠjoki tie 272 mts | 61.64659 | 21.68996 |
| 9 | 280 | 0 | 1.1 | 0.016 | 1 | 1.3 | 1300 | 0.34 | 2.1 | 2.9 | 0 | 0 | 0 | 6/13/13 | 66209::EtelŠjoki tie 272 mts | 61.64659 | 21.68996 |
| 10 | 180 | 0 | 1.2 | 0.011 | 0.42 | 1.3 | 1800 | 0.53 | 1.1 | 1.4 | 0 | 0 | 0 | 7/22/13 | 66209::EtelŠjoki tie 272 mts | 61.64659 | 21.68996 |
| 11 | 170 | 0 | 0.98 | 0.01 | 0.36 | 1.4 | 1500 | 0.45 | 2.5 | 1.5 | 0 | 0 | 0 | 8/7/13 | 66209::EtelŠjoki tie 272 mts | 61.64659 | 21.68996 |
| 12 | 110 | 0 | 0.82 | 0.015 | 0.38 | 1.3 | 1100 | 0.31 | 1.8 | 1.2 | 0 | 0 | 0 | 9/18/13 | 66209::EtelŠjoki tie 272 mts | 61.64659 | 21.68996 |
| 13 | 720 | 0 | 0.74 | 0.047 | 0.7 | 2 | 1400 | 0.45 | 4.7 |  | 0 | 0 | 0 | 11/18/13 | 66209::EtelŠjoki tie 272 mts | 61.64659 | 21.68996 |
| 14 | 740 | 15 | 0.73 | 0.056 | 0.84 | 2 | 1500 | 0.96 | 4.3 | 11 | 0 | 15 | 1.5 | 1/8/14 | 66209::EtelŠjoki tie 272 mts | 61.64659 | 21.68996 |
| 15 | 510 | 0 | 0.7 | 0.051 | 0.76 | 1.7 | 1700 | 0.85 | 4.5 | 9 | 0 | 0 | 0 | 2/11/14 | 66209::EtelŠjoki tie 272 mts | 61.64659 | 21.68996 |
| 16 | 750 | 16 | 0.65 | 0.04 | 1 | 1.7 | 1300 | 0.37 | 4.5 | 9.8 | 0 | 26 | 1.6 | 3/4/14 | 66209::EtelŠjoki tie 272 mts | 61.64659 | 21.68996 |
| 17 | 600 | 0 | 0.65 | 0.027 | 0.85 | 2.1 | 1400 | 0.37 | 4.2 | 11 | 0 | 0 | 0 | 3/31/14 | 66209::EtelŠjoki tie 272 mts | 61.64659 | 21.68996 |
| 18 | 460 | 15 | 0.64 | 0.025 | 1.1 | 1.5 | 1300 | 0.47 | 5.1 | 6.9 | 0 | 26 | 1.5 | 4/8/14 | 66209::EtelŠjoki tie 272 mts | 61.64659 | 21.68996 |
| 19 | 870 | 15 | 0.7 | 0.037 | 1.6 | 1.7 | 1800 | 0.79 | 4.4 | 8.8 | 0 | 35 | 1.8 | 4/15/14 | 66209::EtelŠjoki tie 272 mts | 61.64659 | 21.68996 |
| 20 | 450 | 0 | 0.75 | 0.025 | 1 | 8 | 1300 | 0.76 | 4 | 4.7 | 0 | 0 | 0 | 5/13/14 | 66209::EtelŠjoki tie 272 mts | 61.64659 | 21.68996 |
| 21 | 200 | 12 | 0.78 | 0.07 | 1 | 1.4 | 1200 | 1.5 | 3.8 | 6.6 | 0 | 13 | 1.1 | 6/9/14 | 66209::EtelŠjoki tie 272 mts | 61.64659 | 21.68996 |
| 22 | 190 | 0 | 1.4 | 0.01 | 1 | 1.5 | 1700 | 0.44 | 2.6 | 1.6 | 0 | 0 | 0 | 8/11/14 | 66209::EtelŠjoki tie 272 mts | 61.64659 | 21.68996 |
| 23 | 140 | 0 | 0.63 | 0.01 | 0.36 | 1.2 | 1200 | 0.37 | 2.1 | 2.2 | 0 | 0 | 0 | 9/29/14 | 66209::EtelŠjoki tie 272 mts | 61.64659 | 21.68996 |
| 24 | 200 | 16 | 0.61 | 0.013 | 1 | 0.9 | 1200 | 0.23 | 2.4 | 0 | 0 | 0 | 0 | 10/20/14 | 66209::EtelŠjoki tie 272 mts | 61.64659 | 21.68996 |
| 25 | 0 | 0 | 0 | 0.01 | 0 | 0.19 | 1.6 | 0 | 13 | 2 | 0.23 | 0 | 0 | 7/29/14 | 6451::Kojo 37 Varvoorinjuopa | 61.48256 | 21.85145 |
| 26 | 0 | 0 | 0 | 0.01 | 0 | 0.22 | 1.5 | 0 | 7.9 | 1.7 | 0.1 | 0 | 0 | 8/6/14 | 6451::Kojo 37 Varvoorinjuopa | 61.48256 | 21.85145 |
| 27 | 0 | 0 | 0 | 0.01 | 0 | 0.32 | 2 | 0 | 7.6 | 2.5 | 0.61 | 0 | 0 | 8/13/14 | 6451::Kojo 37 Varvoorinjuopa | 61.48256 | 21.85145 |
| 28 | 0 | 0 | 0 | 0.01 | 0 | 0.28 | 1.5 | 0 | 6.4 | 2.7 | 0.67 | 0 | 0 | 8/19/14 | 6451::Kojo 37 Varvoorinjuopa | 61.48256 | 21.85145 |
| 29 | 0 | 0 | 0 | 0.01 | 0 | 0.18 | 1.3 | 0 | 5.7 | 0.74 | 0.29 | 0 | 0 | 8/27/14 | 6451::Kojo 37 Varvoorinjuopa | 61.48256 | 21.85145 |
| 30 | 0 | 0 | 0 | 0.01 | 0 | 0.32 | 1.9 | 0 | 5.6 | 2.4 | 0.41 | 0 | 0 | 9/10/14 | 6451::Kojo 37 Varvoorinjuopa | 61.48256 | 21.85145 |
| 31 | 0 | 0 | 0 | 0.01 | 0 | 0.37 | 2 | 0 | 5.3 | 9.1 | 0.36 | 0 | 0 | 10/13/14 | 6451::Kojo 37 Varvoorinjuopa | 61.48256 | 21.85145 |

Figure 3: The data structure of water quality in the database.

After we prepared the dataset included spatiotemporal metadata and we create the database of water quality in parts of heavy metal by using PostgreSQL for creating multi-dimensional semantic space to analyze water quality in details.

## 2.3 Equation and Analysis

In this step, we evaluate the historical change in the water-body of Pori's water resources and realize the part of processing (P) on rSPA processes by using the multi-heavy metal parameters to create the index for analyzing the water quality in water resources in a river. The system architecture has 3 parts shown in Figure 4. The first procedure of processes performs the processing part (P) of rSPA processes with index creation (rHMEI) by using multi-heavy metal parameters. Second procedure of processes analyzes data by using an index on multi-dimensional subspace.

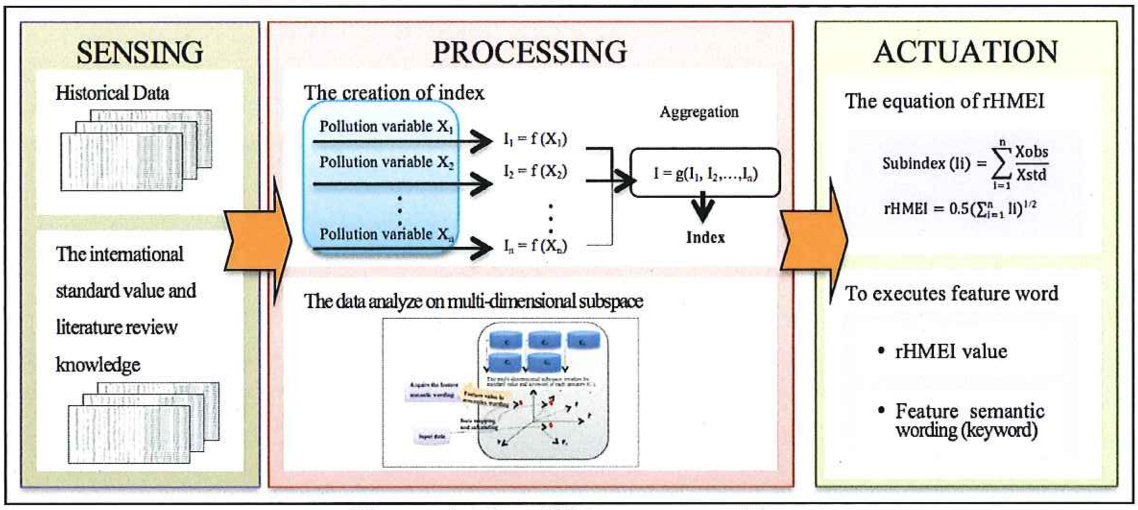


Figure 4: The rSPA system architecture.

### 2.3.1 The procedure of index creation (rHMEI) by using multi-heavy metal parameters

The first procedure of processes of this study, we study the pollutant-environmental variable and environmental indicator for environmental monitoring, public information and scientific research. The environmental indicator is summarized data from the amount of information into the simplest form which decreasing some information data and do not deviate the measured value. We apply and realize the environmental indicator on an increasing scale, which refers to high index where pollutant is high, to be the environmental evaluation indices. The step to create as below:

**The first step**, we study the characteristic of heavy metal as the toxicity and availability in the hydrologic environment and design several parameters for index creation.

**The second step**, we implement the index that consists of sub-indices and the environmental indicator of heavy metal. The summaries of each sub-index as shown in figure 5.

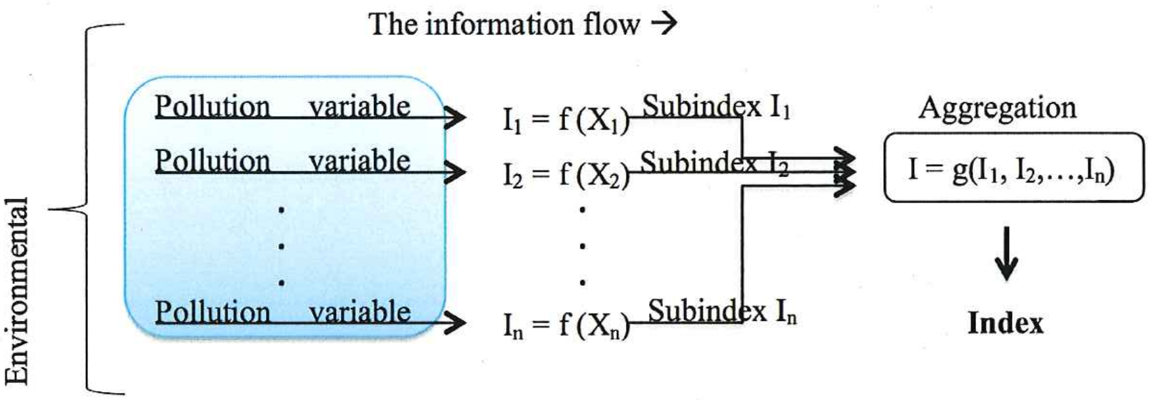


Figure 5: The information flow in rHMEI creation.

*Corresponding author (C. Veesommai). Tel/Fax: +81 80 8863 9389. E-mail addresses: chalisav@sfc.keio.ac.jp, chali-h@hotmail.com  eISSN:
Online Available at

The subindex is calculated using following formulas

$$I_i = f_i (X_i) \qquad (1),$$

When $X_i$ is the variable of each heavy metal substance

$I_i$ is subindex

**The third step**, we implement the relation of sub-index with the variable of substance. The heavy metal is a hazard substance and causes acute effects in the concentration of a substance over the threshold value. The dose response curve of substance's concentration characteristic is segmented in the linear function, and the damage functions of the heavy metal are shown in Figures 6 (a) and (b).

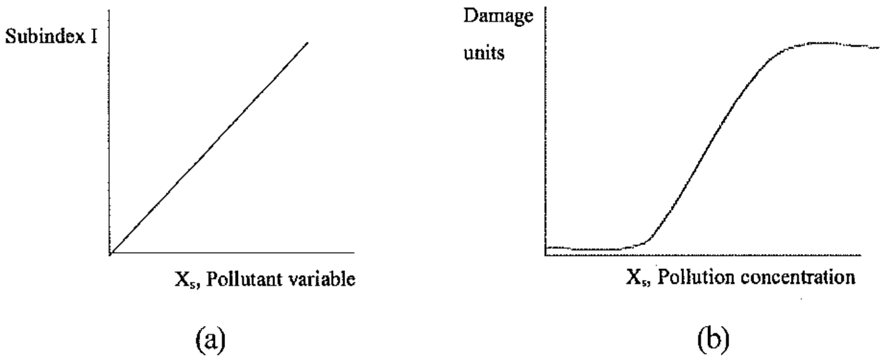


Figure 6: The dose response curve of (a) substance's concentration characteristic (segmented linear function) and (b) the damage function of the heavy metal.

After we study several methods for aggregation of sub-indices and we use the root mean square for eliminating ambiguous and eclipsing situations. The characteristic curve for the root sum power and the root mean square for reducing ambiguous and eclipsing situations shown in Figure 7 (a) and (b).

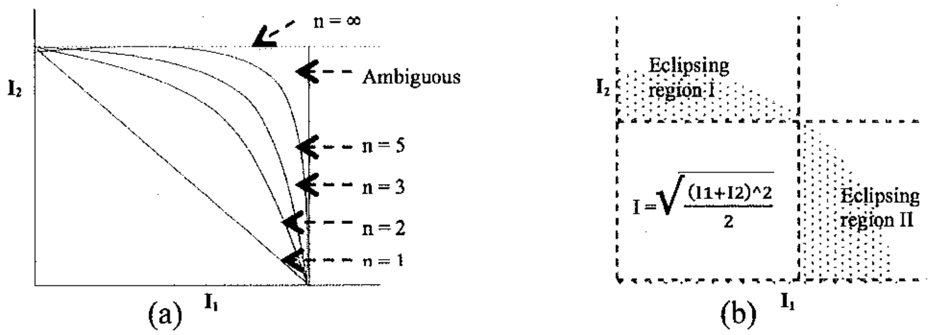


Figure 7: The subindex summarization type of (a) root sum power and (b) root mean square efficiency.

## 2.3.2 The procedure of data analyses by rHMEI on rSPA processes

As second procedures used for analyzing data on rSPA processes, we apply the rHMEI in a processing part on rSPA processes by using postgreSQL shown in Figure 8. The step to processing is explained below:

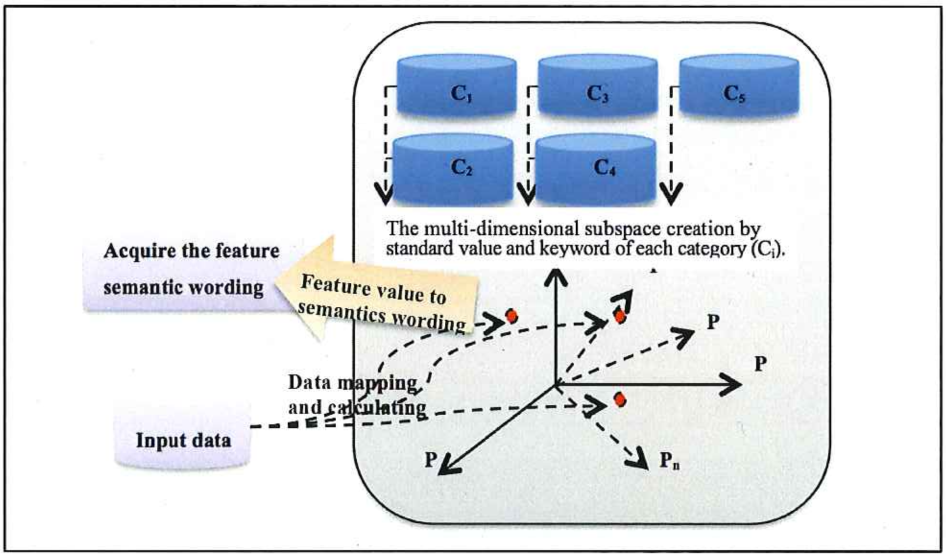


Figure 8: The second procedures used for analyzing data on rSPA processes.

**The first step**, the process for creation of 5 categories as an aquatic life ($C_1$), irrigation ($C_2$), estuary and harbor basin water ($C_3$), livestock ad wildlife ($C_4$), and industrialization ($C_5$) by multi-heavy metal-parameters and the creation of wording based on standard threshold toxic value class in the database as an effective wording such as safe, threshold toxic, excellent, hazard, optimum, damage, satisfactory and unfit.

**The second step,** the process for selecting the multi-heavy metal-parameters that relate to the category group and creating the multi-dimensional semantic space.

**The third step,** This process is for mapping input data and calculating the rHMEI by using the multi-dimensional semantic space of each category.

**The fourth step,** This process is for executing feature word processing by selecting the candidate important word in the range of parameter in the heavy metal quality field.

## 3 Result and Discussion

In terms of the specific characteristic of the heavy metal parameter and several methods for the aggregation of sub-index, we apply the root mean square for creating the index to analyze the water quality in water resources in the processing (P) part of rSPA

*Corresponding author (C. Veesommai). Tel/Fax: +81 80 8863 9389. E-mail addresses: chalisav@sfc.keio.ac.jp, chali-h@hotmail.com  eISSN:
Online Available at

processes. The results of first and second procedures are described below

## 3.1 The processes of index creation (rHMEI) by using multi-heavy metal parameters

### 3.1.1 The heavy metal parameter design for index creation

We use the low to the high toxicity of heavy metal parameters (9 parameters; Arsenic, Cadmium, Chromium, Copper, Iron, Lead, Mercury, Nickel and Zinc) as U.S. Geological Circular 1133 (Robert H.M., 1995) for creating the river Heavy Metal Evaluation Index (rHMEI). The specific characteristics of the heavy metal parameter for index creation are shown in Table 2 (Duruibe J.O. *et al.*, 2007, Mung D. *et al.*, 2014, Baigal A.T. *et al.*, 2015, Jin H.K. *et al.*, 2010, Thanh k. *et al.*, 2015, John B.W. *et al.*, 2015 and Hongxia S. *et al.*, 2016).

Table 2: The heavy metal parameter design for index creation.

| Parameter | Description |
|---|---|
| Arsenic (As) | Arsenic is the most toxicity substances and appears in tree allotropic form. In the environment, arsenic can be found naturally on the earth in small concentration that occurs in soil and minerals and then enters the water, and human activities such as mining, melting and copper-lead-zinc producing industries. In the environmental effects of arsenic in water is chronic toxicity for a time period in aquatic life, carcinogenic arsenic (III) compound which blocks enzymatic processes, increasing its toxicity in an animal. Exposure to inorganic arsenic can cause various health effects such as skin disturbances, declined resistance to infections and damage DNA. Exposure organic arsenic can cause cancer, DNA damage. |
| Cadmium (Cd) | Cadmium is a high toxicity metal as low concentration exposure and environmental pollutant classified as a carcinogen group 1. Cadmium can found in several industrial activities as paints, manufacturing of batteries and agricultural industry. In the environmental effects, Exposure to a high value can cause cancer. |
| Chromium (Cr) | Chromium is one of the toxic metal in water resource areas. In the environment, chromium can found in the earth's crust and human activities as a manufacturing. In the environmental effects, chromium can be transported and absorbed by sludge when high concentration can be extremely dangerous to aquatic life and vegetation. Exposure to a high value can cause skin disturbances, accumulated in kidneys. |
| Copper (Cu) | Copper is one of 129 priority pollutants in listed by EPA. In the environment, copper can found in the earth's crust and human activities as a manufacturing. In the environmental effects, copper can be extremely dangerous and to aquatic life and vegetation when dissolved in the water. Exposure to a high value can cause acute- chronic health effects, cancer hazard, and reproductive hazard. |

Table 2: The heavy metal parameter design for index creation (*Cont.)*

| Parameter | Description |
|---|---|
| Iron (Fe) | Iron is one of heavy metal in water resource areas. In the environment, chromium can found in the earth's crust and human activities as a manufacturing. In the environmental effects, Iron (III)-O-arsenite, pentahydrate can be hazardous to the environmental, special attention should be given to plants and water. Exposure to a high value can cause conjunctivitis, choroiditis, and retinitis if it contacts and remains in the tissues. |
| Lead (Pb) | Lead is one out of four metals that have the most hazarding effects on human health but can cause several effects as a serious damage. In the environment, lead occurs naturally and mainly from human activities such as car engines burned. In the environmental effects of lead in water is accumulates in the aquatic organisms and soil, small concentration can cause effects to shellfish, fish, and phytoplankton. Exposure to lead can cause a rise in blood pressure and disruption of the biosynthesis. |
| Mercury (Hg) | Mercury is one of molecular that widespread environment toxicant and pollutant which severe alternations in body tissue and also cause a wide range of human effect. Exposure to mercury is possible human carcinogens and determined by EPA as damage to nervous system, nervous system is very sensitive to all of the form of mercury, damage to brain function, kidneys and developing features in high concentration of mercury, damage to lung, nausea, vomiting, diarrhea, and blood (blood pressure increase) or heart rate, skin rashes and eye irritation as high level in short-term exposure. |
| Nickel (Ni) | Nickel is one of four substances that are ferromagnetic at room temperature (25 $^0$C). In the environment, nickel found the small amount and occurs by combined with sulfur. In the environmental effects of high nickel concentration in water is can diminish the growth rates of algae, hazard to microorganisms and aquatic plant. Exposure to nickel can cause skin damage, allergic reactions, destroy the development of organisms. |
| Zinc (Zn) | Zinc is a fairly active element. It dissolves in both acids and alkalis. In the environment, zinc occurs naturally and mainly from human activities such as car engines burned. In the environmental effects of lead in water is accumulates in the aquatic organisms and soil, small concentration can cause effects to shellfish, fish, and phytoplankton. Exposure to zinc can cause experience loss of hair, skin lesions, skin rashes and sore throat. |

From Table 2, the difference toxic substances of heavy metals are occurred naturally and produced by human activities and make effects for a living organism and persistent environmental contaminants as stable oxidation states.

*Corresponding author (C. Veesommai). Tel/Fax: +81 80 8863 9389. E-mail addresses: chalisav@sfc.keio.ac.jp, chali-h@hotmail.com  

### 3.1.2 The implementation of index that consists of sub index

The rHMEI is a mathematics instrument used to aggregate diverse heavy metal parameters and their multi-dimensional aspects into a single score. The equation for calculating the rHMEI is defined in the following:

$$\text{Subindex (Ii)} = \sum_{i=1}^{n} \frac{\text{Xobs}}{\text{Xstd}} \quad (2),$$

$$\text{rHMEI} = 0.5\left(\sum_{i=1}^{n} \text{Ii}\right)^{1/2} \quad (3),$$

When $X_{obs}$ is a concentration of each heavy metal

$X_{std}$ is a threshold of each heavy metal parameter in each category (Aquatic life, livestock and wildlife, Irrigation, Industrial and Estuary Basic water)

The rHMEI classification into 2 classes based on the international standard and/or maximum values are shown in Figure 9.

Output pane

Data Output | Explain | Messages | History

| | id integer | category_id integer | range_lower numeric | range_upper numeric | keyword character varying |
|---|---|---|---|---|---|
| 1 | 1 | 1 | 0.000 | 2.13 | Safe for AquaticLife |
| 2 | 2 | 1 | 2.13 | 1000.000 | Threshold toxic for AquaticLife |
| 3 | 3 | 2 | 0.000 | 2.13 | Excellent for Irrigation |
| 4 | 4 | 2 | 2.13 | 1000.000 | Hazard for Irrigation |
| 5 | 5 | 3 | 0.000 | 2.13 | Optimum for Estuary and HarbourBasinWater |
| 6 | 6 | 3 | 2.13 | 1000.000 | Damage for Estuary and HarbourBasinWater |
| 7 | 7 | 4 | 0.000 | 2.13 | Satisfactory fro Livestock and poultry |
| 8 | 8 | 4 | 2.13 | 1000.000 | Threshold toxic for Livestock and poultry |
| 9 | 9 | 5 | 0.000 | 2.13 | Optimum for Industrial process |
| 10 | 10 | 5 | 2.13 | 1000.000 | Unfit, high corrosive, scaling and fouling for Industrial process |

Figure 9: The rHMEI classification of each category.

We compare the mathematics instrument of rHMEI with several mathematics instruments in the water quality field shown in Table 3.

Table 3: The comparison of several methods and our method.

| Aggregation function | Relate research | Increasing scale indices | Decreasing scale indices |
|---|---|---|---|
| Weighted linear sum | - Water Quality Index (WQI) (Fabiano D.S.S. *et al*, 2008) | Eclising and no ambiguity | Eclising and no ambiguity |
| | - Water Quality Index (WQI) (Summiya N. *et al*, 2014) | | |
| | - Heavy metal pollution Index (HPI) (Mona H.M. *et al*, 2014) | | |
| Linear sum | Metal Index (MI) (Mohamed E.G. *et al*, 2014) | No eclising and ambiguity | Eclising and no ambiguity |
| | Pollution Index (PI) (Mohamed E.G. *et al*, 2014) | | |
| Root sum power | Polluton load Index (PLI) (Amirhossein P. *et al.*, 2015) | Eclising and no ambiguity | Eclising and no ambiguity |
| Root Mean square | Our current study | Minimized eclising and ambiguity as n approch ∞ | Eclising and no ambiguity |

After the implementation of the index that consists of sub-indexes and compares the mathematical instrument of rHMEI with several mathematical instruments in the water quality field. We found that our current study makes an advantageous effect because it can decrease the eclipsing region which is a problem in the case of a nonlinear function making the resulting error values from an ideal situation.

## 3.2 The data analyze by rHMEI on rSPA processes

We apply the rHMEI as a processing part on rSPA processes by using postgreSQL. The rHMEI for Aquatic life, livestock and wildlife, Irrigation, Industrial, and Estuary Basic water category was computed using the guidelines of the standard in the water quality of water resource in each category (FAO, 1985), (CCME, 2007), (UNECE, 1994), (WHO, 1989) and (Ministry of industrial, 1978).

In the processing part on rSPA processes using PostgresSQL, the system computes the rHMEI value and detects the critical situation by using the standard threshold value of each category. The result of feature wording as an actuation on rSPA found that the process will execute feature word processing by detecting rHMEI value in the range of parameters in heavy metal of each category:

- The feature word of Aquatic life found safe for aquatic life 128 notifications and

*Corresponding author (C. Veesommai). Tel/Fax: +81 80 8863 9389. E-mail addresses: chalisav@sfc.keio.ac.jp, chali-h@hotmail.com  eISSN:
Online Available at

threshold toxic for aquatic life 43 notifications.

- The feature word of Irrigation found excellent for irrigation 142 notifications and hazard for irrigation 29 notifications.
- The feature word of Estuary and harbor basin water found all of the notifications in Pori's water resource shown Optimum for estuary and harbor basin water.
- The feature word of Livestock and wildlife found satisfactory for livestock and poultry 149 notifications and threshold toxic for livestock and poultry 22 notifications.
- The feature word of Industrial found the optimum for industrial process 170 notifications and Unfit, high corrosive, scaling and fouling for Industrial process 1 notifications.

The results in detail are shown in Appendix 1.

## 4 Conclusion

This paper has presented the mathematical instrument as an index for developing the meta-level knowledge in the database system for water resources with the method for classification and interpretation of monitoring data on rSPA processes. The significant index is the river Heavy Metal Evaluation Index (rHMEI), expressed in the dimensional subspace of heavy metal in Pori's water resource, Finland.

The significant result of feature-wording as an actuation on rSPA as A spot (Eteljoki tie 272 mt) is the threshold toxic for aquatic life and hazard for irrigation. The threshold toxic is also the result for livestock and poultry, during 4 seasons of 2013-2014. Furthermore, V spot (Rahkakeitaan kp kaiv) is the threshold toxic for aquatic life during the summer season of 2014-2015, and X spot (Rahkakeitaan kp oja it) is the threshold toxic for aquatic life during the summer season of 2013-2015. Totally the spot of Pori's water resources is suitable for estuary and harbor basin water and industrial process.

The rHMEI is a flexible tool to calculate and classify water resource quality in a heavy metals substance within a simple framework, and it is effective for increasing scale of indices (heavy metals substance) while minimizing the eclipsing and ambiguity in the deep water quality analysis.

As a future work, our current method will be extended to apply it to anthropogenic addition of nutrient compounds that lead to the accumulation in water bodies which causes eutrophication crisis. This extention and application will make the progress in classifying the eutrophication situations with the latest deep learning method, and also in evaluating the system with symbolic-filtering and semantic-ordering functions for the analysis of water quality in wide areas.

## 5 Acknowledgements

This work is supported partially by MEXT Grant-in-Aid for the Program for Leading Graduate School Graduate School of Media and Governance, Keio University, and Taikichiro Mori Memorial Research Funding.

## 6 References

Adam M.S., Elizabeth M.T., Patricia A.F., Hui C.W.S., Marilyn M. Gregory D.B. and John R. (2015). Mercury concentrations in Atlantic bottlenose dolphins (Tursiops truncatus) inhabiting the Indian River Lagoon, Florida: Patterns of spatial and temporal distribution. *Marine Pollution Bulletin*, 97, 544-547.

Amirhossein P., Gholamreza N.B., Mojtaba A., Mohsen S. and Akbar N. (2015). A new index for assessing heavy metals contamination in sediments: A case study. *Ecological Indicator,* 58, 365-373.

Baigal A.T., Anna M.R. and Helena G. (2015). Arsenic toxicity effects on microbial communities and nutrient cycling in indoor experimental channels mimicking a fluvial system. *Aquatic Toxicology*, 166 (2015), 72-82.

Canadian Council of Ministers of the Environment. (2007). Canadian environmental quality guidelines. Canada.

Das A.K. (1990). Metal ion induced toxicity and detoxification by chelation therapy: A text book on medical aspects of bio-inorganic chemistry, Ist ed, CBS, Delhi, p. 17-58.

Duruibe J.O., Ogwuegbu M.O. and Egwurugwu J.N. (2007). Heavy metal pollution and human biotoxic effects. *International Journal of Physical Science*, 2 (5), 112-118.

Fabiano D.S.S., Altair B.M., Marcia C.B., Sonia M.N.G. and Maria J.S.Y. (2008). Water quality index as a simple indicator of aquaculture effect on aquatic bodies. *Ecological Indeicators,* 8, 476-484.

FAO. (1985). Food and Agriculture Organization: Water Quality for Agriculture. *Irrigation and Drainage Paper*, 29 (1). United Nations, Rome.

Finnish Environment Institute. (2016). Finnish Environment Institute SYKE, Metadata Portal -Surface water quality. Retrieved from http://metatieto.ymparisto.fi:8080/geoportal /catalog/search/resource/details.page?uuid=%7BB1444E19-0F36-49F5-A849-01A3D

*Corresponding author (C. Veesommai). Tel/Fax: +81 80 8863 9389. E-mail addresses: chalisav@sfc.keio.ac.jp, chali-h@hotmail.com  eISSN:
Online Available at

2083A11%7D

Hongxia S., Wenjing W., Jixing G., Rong X., Fengze J., Lingyan Z. and Guren Z. (2016). Effects of nickel exposure on testicular function, oxidative stress, and male reproductive dysfunction in Spodoptera litura Fabricius. *Chemophere,* 148, 178-187.

Jin H.K., Jae S.R., Jin S.L., Hans U.D. and Jehee L. (2010). Effect of cadmium exposure on expression of antioxidant gene transcripts in the river pufferfish, Takifugu obscurus (Tetraodontiformes). *Comparative Biochemistry and Physiology,* Part C (152), 473-479.

Jonh B.W., James M.L., Anthony R.O. and Angela B. (2015). A national statistical survey assessment of mercury concentrations in fillets of fish collected in the U.S. EPA national rivers and streams assessment of the continental USA. *Chemophere,* 122, 52-61.

Kiyoki Y. and Chen X. 2015. Contextual and Difference Computing for the Multi-Dimensional World Map with Context-Specific Spatial-Temporal and Semantic Axes. *Information Modeling and Knowledge Bases XXV,* 82-97.

Kiyoki Y. and Kawamoto M. (2007). Semantic Associative Search and space Integration Methods Applied to semantic Metrics for Multiple Medical Fields. *Information Modelling and Knowledge Bases XVIII,* 120-135.

Kiyoki Y., Kitagawa T. and Hayama T. 1994. A Metadatabase System for Semantic Image search by a Mathematical Model of Meaning. *ACM SIGMOD Record*, 23 (4), 34-41.

Kiyoki Y., Sasaki S. and Koopipat C. (2015). Multi-Dimension Semantic Computing with Spatial-Temporal and Semantic Axes for Multi-spectrum Images in Environment Analysis. *Information Modeling and Knowledge Bases XXV*, 82-97.

Open Knowledge International. (2016). Open Knowledge Foundation, The Open Definition. Retrieved from http://opendefinition.org/

Meng D., Dongbin W., Zhuowei T. Aiwu L. and Yuguo D. (2014). The potential risk assessment for different arsenic species in the aquatic environment. *Journal of Environmenta Sciences,* 27, 1-8.

Mohamed E.G., Ali M.H., Ibrahim A.A., Ayman H.F. and Seliem M.E. (2014). Evaluation of surface water quality and heavy metal indices of Ismaila Canal, Nile River, Egypt. *Egyptian Journal of Aqualtic Research,* 40, 225-233.

Mona H.M., Gholam R.L. and Pooria D. (2014). Assessing the effect of heavy metal concentrations (Fe, Pb, Zn, Ni, Cd As, Cu, Cr) no the quality of adjacent ground water resource of knorasan steel complex. *International Journal of Plant, Animal and Environmental Science,* 4 (2), 511-518.

Robert H.M. (1995). Contaminants in the Mississippi River 1987-92. U.S Geological Survey Cicular *1133*, 55.

Summiya N., Muhammad Z.H. and Riffat N.M. (2014). Heavy metals disbution, risk assessment and water quality characterization by water quality index of River Soan, Pakistan. *Ecological Indicator,* 43, 262-270.

Thanh K.L.B., Chi D.H., Thanh S.D. and Tham C.H. (2015). Copper toxicity and the influence of water quality of Dongnai River and Mekong River waters on copper bioavailability and toxicity to three tropical species. *Chemophere,* 144, 872-878.

UNECE. (1994). Standard Statistical Classification of Surface Freshwater Quality for the Maintenance of Aquatic Life. In: Readings in International Enviroment Statics, united Nations Economic Commission for Europe, United Nations, New York and Geneva.

Veesommai C., Kiyoki Y., Sasaki S. and Chawakitchareon P. (2016). Wide-Area River-Water Quality Analysis and Visualization with 5D World Map System. *Information Modelling and Knowledge Base XXVII*, 280, 31-43.

Veesommai C. and Kiyoki Y. (2015). River Water-quality Analysis: Critical Contaminate Detection, Classification of Multiple-water-quality-parameters Values and Real-time Notification by rSPA. *IEEE Press*, 206-211.

Veesommai C., Kiyoki Y. and Sasaki S. (2016). A Multi-dimensional River-water Quality Analysis System for interpreting Environmental Situations. *Proceedings of the 25th International Conference on Information Modelling and Knowledge Bases, Tampere, Finland*, 1, 48-68.

WHO. (1989). World Health Organization: Health Guidelines for the Use of Wastewater in Agriculture and Aquaculture. *Report of a Scientific Group Meeting. Technical Report Series*, no 778. Geneva.

---

Y. Kiyoki is a Professor of Faculty of Environment and Information Studies at Keio University, Japan. He received his B.E., M.E. and Ph.D. degrees in electrical engineering from Keio University in 1978, 1980 and 1983, respectively. In 1983 and 1984, he worked at Electrical and Communication Laboratory of NTT. From 1984 to 1996, he was with Institute of Information Sciences and Electronics, University of Tsukuba, as an assistant professor and then an associate professor. Since 1996, he has been with Department of Environmental Information at Keio University, where he is currently a professor. His research addresses multidatabase systems, knowledge base systems, semantic associative processing, and multimedia database systems. He serves as the chair on Special Interest Group of Database Systems and as the editor-in-chief on Transactions on Databases in Information Processing Society of Japan. He also serves as the program chair for the 7th International Conference on Database Systems for Advanced Applications.

H. Jaakkola is a Professor in Software Engineering in the Tampere University of Technology Finland. He obtained his MSc (Computer Science, 1974), Licentiate degree (1982, Computer Science) and BSc (1982, Business Economics) in the University of Tampere, and PhD (1990, Electrical Engineering) in the Tampere University of Technology - TUT). Since 1974, he was working in the variety of teaching positions in the University of Tampere. In 1984, he was moving to the Tampere University of Technology, first for the Professor position and in 1985 he was nominated to the director of the Research Institute for Information Technology (TUT). In 1992, he was invited to the position of professor (the Industrial Applications of ICT; Software Engineering) in the Pori Department of TUT. In this position, he has been Director of TUT Pori Campus from 1992 to 2007. Currently, he works as full time Research Professor in the same Institute. Hannu Jaakkola has the Title of Docent in the University of Lapland (since 1998) and he has been working as a visiting professor in Keio University Japan in 2015 (Keio Super Global Program). He is also the member of the International Advisory Board of the GESL Program in Keio University. The research work of Hannu Jaakkola covers a variety of ICT related topics, currently focused on the multicultural distributed software engineering. Other current activities relate to information society phenomena, from technical aspects to social networking, including context recognition and the use of open and big data technologies. He has been the director of six Executive MBA Programs (since since 2003). He is member of the Association for Computing Machinery (ACM) and the Institute of Electrical and Electronics Engineers (IEEE). He has published more than 300 scientific publications. Hannu Jaakkola has wide variety of duties in international academic

*Corresponding author (C. Veesommai). Tel/Fax: +81 80 8863 9389. E-mail addresses: chalisav@sfc.keio.ac.jp, chali-h@hotmail.com  eISSN:
Online Available at

community (conference committees, journal reviewer, academic assessments, etc.) and has awarded of his work by several organizations. In this sector of activities it is worth to mention his role as one of the founders of the International Conference on Information Modelling and Knowledge Bases (formerly European Japanese – EJC), which has been organized annually since 1988 and plays an important role in collaboration between Japanese and Finnish researchers.

J. Soini (Dr.Tech) has a permanent position as research manager at Tampere University of Technology (TUT) in Pori department, Finland. Since 2004 he has been working in the academic sector, and his study area relates to software engineering. Previously he worked in the private sector as a software expert and software consultant for Innomat Ltd and i2 Technologies Finland Ltd. Currently his duties include to lead the research carried out in the Software Engineering and Intelligent Systems (SEIntS) research group and also preparing, leading and managing software engineering research projects in the Pori department. His recent research has been focused on wide databases, complex data network structures (both fixed and mobile) and the services built on them.

P. Chawakitchareon is an Associate Professor of Environmental Engineering Department at Chulalongkorn University, Bangkok, Thailand. She received her B.Sc and M.Sc from Mahidol University, Bangkok, Thailand and Ph.D. degrees in Environmental Engineering from ENTPE-Lyon I, France in 1982, 1984 and 1990 respectively. Her current research interests involve utilization of industrial waste for environmental engineering applications and environmental water quality monitoring.

P. Sillberg received M.Sc. from Tampere University of Technology (TUT), Finland, in 2009. He is working as a researcher in the TUT Pori Department in Pori, Finland. His current research interests are in context sensitive user interfaces, mobile device programming and web development.

C. Veesommai is a Ph.D. candidate of Global Environmental System Leader Program (GESL), Graduate School of Media and Governance, Keio University, Japan. She received her B.Sc. and M.Eng. from Chulalongkorn University, Bangkok, Thailand in 2010 and 2012, respectively. Miss Veesommai current interests involve application of river water-quality analysis and an application of utilization of industrial waste.