# An e-Process Selection Meta-Model using four Decision Making processes

**Frina Albertyn**

Eastern Institute of Technology
New Zealand
*falbertyn@eit.ac.nz*

*Abstract*: **This journal paper introduces four different decision making methods, their application on e-Process Selection and then the requirements of a meta-model for each that will support the e-Process Selection Methodology. E-Processes are those development processes that can be used to develop e-Commerce Information Systems (eCIS). The four different decision making processes used to build the metal models are Value-Benefit Analysis, Analytical Hierarchy Process, Case-Based Reasoning and a Social Choice Method. This journal paper then uses each of these meta-models to develop an integrated meta-model for e-Process selection.**

*Keywords*: Decision making, development processes, e-Commerce, e-Process.

## I. Introduction

The key objective of this research was to develop a selection methodology that can be used to support and aid the selection of development processes for e-Commerce Information Systems (eCIS) effectively using various decision methods. The selection methodology developed here supports developers in their choice of an e-Commerce Information System Development Process (e-Process) by providing them with a few different decision making methods for choosing between defined e-Processes using a set of quality aspects to compare and evaluate the different options. The methodology also provides historical data of previous selections that can be used to further support their specific choice.

The research was initiated by the fast growing Information Technology environment, where e-Commerce Information Systems is a relatively new development area and developers of these systems may be using new development methods and have difficulty deciding on the best suited process to use when developing new eCIS [45]. These developers also need documentary support for their choices and this research helps them with these decision-making processes.

The e-Process Selection Methodology allows for the comparison of existing development processes as well as the comparison of processes as defined by the developers. Four different decision making methods, the Value-Benefit Method (Weighted Scoring), the Analytical Hierarchy Process, Case-Based Reasoning and a Social Choice method are used to solve the problem of selecting among e-Commerce Development Methodologies.

Decision making in the business environment needs to be of a high standard. This is a difficult and complex process. There are different components required for any specific solution to a problem. These components include things such as technology, material, development and human resources which tend to be expensive. The decision maker is responsible for deciding which solution will be selected and this decision can potentially be very good or potentially be a risky exercise.

Today's market is very competitive, which means it is important to make decisions which is based on well-motivated reasons, which provides the necessary effect and are taken quickly [26], [30], [43]. The decision maker most probably require some support in this decision making process.

Decision making can be described as the cognitive process leading to the selection of a course of action among a set of possible candidate solutions. Every decision making process produces a final choice. It can be an action or an opinion. It begins when we need to do something but do not know what. Therefore, decision making is a reasoning process which can be rational or irrational; it can be based on explicit assumptions or tacit assumptions. In Doumpos & Zopounidis [15] multi-Criteria Decision Analysis (MCDA), or Multi Criteria Decision Making (MCDM), is defined as being a discipline aimed at supporting decision makers who are faced with making numerous and conflicting evaluations. MCDA aims at highlighting these conflicts and deriving a way to come to a compromise in a transparent process.

In this journal paper four different MCDM decision processes (Value Benefit Method (VBA), the Analytic Hierarchy Process (AHP), Case-Based Reasoning (CBR) and Social Choice Methods (SCM)) that have been researched and a meta-model developed for each. These meta-models are then used to develop an integrated meta-model for e-Process selection (this is development processes that can be used to develop e-Commerce Information Systems).

When you submit your paper print it in two-column format, including figures and tables. In addition, designate one author as the "corresponding author". This is the author to whom proofs of the paper will be sent. Proofs are sent to the corresponding author only.

## II. Applying the Decision making Methods

### A. *Value Benefit Analysis (Weighted score) (VBA)*

VBA is a method that allows the selection of one option, in our case one e-Process, from a set of options based on an evaluation process. This procedure does not focus on cost as a criterion, but rather specifies the identification of a number of criteria that can be used for evaluating and comparing the different options against one another. The application of Value Benefit Analysis is also called the weighted scoring method.

Scoring methods were developed originally to overcome some of the biases of other selection methods. A simple scoring method developed is the un-weighted 0-1 factor method, which lists multiple criteria of significant interest to the people making the decision on what to select. The selection committee is given a list of the chosen criteria and a portfolio of the each of the potential projects. This committee then checks off those criteria that the specific project fulfils. The "winning" project is the project that has the largest number of check marks [24]. In the selection process it was, however, found that often not all criteria are of equal importance. Various projects may satisfy each criterion to different levels. This led to the development of a more sophisticated method of selection called the weighted scoring method [12], [24]. The weighted scoring method is a disciplined subjective approach, using multiple criteria, in selecting a rationally favoured choice from a field of candidate choices [12].

The mathematical formula is:
Sum for a specific project is:

$$Sj = \sum_{i=1}^{h} Sij\,Wj$$

With sij = the score of the ith project on the jth criterion, wj = the weight of the jth criterion, i = 1 to the number of projects and j = 1 to number of criterion.

In this method a set of C = {c1 ... cn} criteria is identified for use to consider and select between different software projects – in our case e-Processes. Too few criteria (< 3) may be too small to work. The importance of each of the criteria is defined in terms of weights (w). The sum of all the weights over all the n criteria is usually set to 1.0 or 10 It is suggested that the number of criteria used in the evaluation be restricted for ease of use and to exclude those criteria that only marginally contribute to the final selection [12], [24].

The value assigned to each weight can be determined in a number of ways. These include the results of surveys, group composite beliefs such as average of values between a group of members and the subjective belief of an individual [24]. In order to choose between the different projects it is also necessary to assign a score sij to each of the projects (i) for each of the criterion (j). Usually a 5 point score will be used [24]. Excellent performance of a specific criterion will produce a top score of five while three is reserved for an average performance and one for a poor performance of that specific criterion. This 5 point scale can be used for qualitative as well as quantitative data thus making it possible to use both criteria which can easily be measured (hard data) as well as soft data such as fit with the organisation's goals, comfort, ease of use and others [24]. Each score is multiplied by its corresponding criterion weight and then for each process these values are summed to determine a total weighted score.

### 1) VBA e-Process ontology

The research contained in this section is based on parts of the following publications; see [4], [5], [6]. Similar to Fettke and Loos [31], Opdahl et al. [32] and Wand et al. [33], following Tom Gruber, the research in this section considers ontology as a specification of a conceptualisation that is shared by a number of people. For the purpose proposed here, i.e., assessing e-Processes, it is best that this specification takes the form of a list of concepts (including definitions of these) that enables the characterising of the e-Processes. A set of 54 quality aspects are used.

Once the associated scales have been assigned to the e-Process quality aspects then e-Processes can be conceptualised as points in a number space, the dimension of which is the number of quality aspects assumed. In order to scale up or drill down respectively, the dimensions in this space are aggregated or decomposed. Considering super- and subspaces in this way allows for the inclusion of further quality aspects if that should be required. Also, quality aspects might be blinded out for particular purposes such as managerial discussions regarding e-Process selection. This means that certain quality aspects may not or must not affect the selection process in which case the weight factors of these quality aspects are simply set to zero.

At the highest level of abstraction high level quality aspects are used to describe e-Processes. With these high level quality aspects the aim is to address e-Process aspects which can be assumed to be important when choosing one of them for a particular development task. Refining the quality aspects, grouping or naming them differently, or adding some of those not considered, may impact a particular choice being made for a given project. However, this would not invalidate this particular selection method.

The e-Process selection methodology associates with each e-Process profile a vector of numbers and for each quality aspect a weighting factor. This will enable the quantitative comparison of the e-Processes and allow the one with largest weighted sum to be selected as the chosen one. To further improve the selection process, weak-point analysis will also be used, see Böhm and Wenger [10] for more detail regarding this method, to analyse different types of improvements to the e-Process selected that would be reasonable with respect to a development task at hand. Both of these methodological ideas, however, presuppose that for each of the quality aspects a numerical scale is available.

### 2) E-Process VBA Selection Methodology

This section defines a meta-heuristic that is based on the idea of patterns. An e-Process pattern here is considered as a

triplet (context, problem, e-Process) where it is assumed that the e-Process is an acceptable solution to the problem in the given context. For each e-Process a score is introduced to measure how well it suits as a problem solution in a given context. The meta-heuristic can be defined as:

```
LOOP
        CHOSE an e-Process from the list
        APPLY the e-Process
        ASSESS its success
        UPDATE the e-Process score
ENDLOOP
```

To formulate this initial heuristics some conventions are introduced. Let there be k high level qualiy aspects d1….dk, m second level quality aspects c1,…,cm and n e-Processes P1, … , Pn. Then each e-Process can be represented by an m-tuple of numbers between 0 and 1.

For each high level quality aspect dj , $j \in \{1, … , m\}$ or second level qualiy aspect ci, $i \in \{1, … , m\}$, and each process Pj, $j \in \{1, … ,n\}$, an enterprise staff, who is an expert in the field, is asked to determine the weight w(1), …, w(m) of the quality aspects c(1), …, c(m), the weight of high level quality aspect d1….dk as well as the performance p(1,j), …, p(m,j) of process Pj with respect to quality aspects c(1), …, c(m) respectively. The numbers w(i) are chosen such that $\sum w(i) = 1$ and $0 \le w(i) \le .1, \forall i \in \{1, …,m\}$. The initial heuristics is then:

1. $C := \varnothing$
2. Chose $J \subseteq \{1, …, n\}$ such that for $j \in J$ the number $S(j) = \sum_{i \in \{1, …, n\}} d(i)*w(i)*p(i,j)$ is maximal and define C := J, observe to chose J maximal. Set $C := C \cup J$.
3. For all $j \in \{1, …, n\} \setminus J, k \in \{1, … , m\}$ perform
   a sensitivity analysis, i.e.,
   Calculate $S(k,j) = \sum_{i \in \{1, …, m\} \setminus \{k\}} w(i)*p(i,j)$.
   Chose sets W, P such that
   $w(k) \in W, p(k, j) \in P$ and determine $T(k,j) = \max\{w*p \mid w \in W, p \in P\}$.
   It will often be convenient to chose W and P such that
   $W = \{cW + hW*\Delta W \mid hW \in \{0, …, rW\}\}$,
   $P = \{cP + hP*\Delta P \mid hP \in \{0, …, rP\}\}$.
   If $S(k, j) + T(k, j) > S(j)$, then the values w,
   p for which the maximum T(k, j) was achieved
   need to be investigated. If these are reasonable
   and acceptable, then redefine $C := C \cup \{k\}$.
4. Do weak point analyses for each candidate in
   $j \in C$, i.e., determine those quality aspects with high
   impact (weight higher than for, e.g. 70 % of the quality
   aspects) and low performance (performance lower than
   for, e.g. 70 % of the quality aspects).
   For each of the weak points consider the performance
   assessment and weight. If one of these should be corrected
   then do so.
   If weak points remain after a. then either $C := C \setminus \{j\}$ or
   replace Pj by an improved version Q scoring no less than
   Pj, and assess it.
5. If the weak point analysis in 4 does not change anything
   then chose among the candidates in C according to a
   predefined strategy. Otherwise go back to 3.

Note that the definition of weak point used here is somewhat arbitrary in that the thresholds of 70 % are not justified. One thus can attempt tuning the heuristics changing these values. In particular the threshold values could be chosen differently. In the next section the Value Benefit Method meta-model is presented.

### 3) E-Process Selection VBA Meta-Model

In Figure 1 and Table 1 the different entities required defining the meta-model for Value-Benefit Analysis Method are stored. The final descriptions are in Table 1. These quality aspects include identification of all the Candidate e-Processes that are going to be used to calculate the "winner" using VBA; capturing all the quality aspects, both the high level aspects as well as the second level aspects.

Each high level quality aspect may have 0 to many second level quality aspects associated with it; the experts (assessors) identify the scores for each of the e-Processes and their associated quality aspects. (This is then semi-permanently stored in the system); the project will, for each of the quality aspects, have a one to many association with the quality scores. After completing the VBA calculations - weak-point analysis and sensitivity analysis are performed on the results to determine whether there might be a better match. This information will also be captured and forms part of the meta-model.
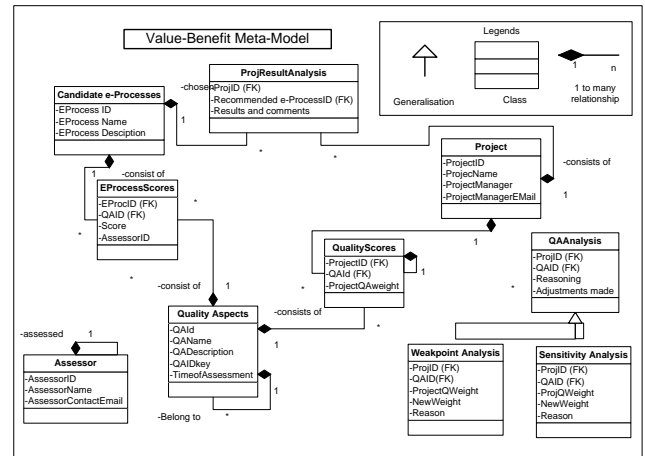


**Figure 1.** e-Process Selection VBA Meta-Model

### B. Analytic Hierarchy Process

The Analytic Hierarchy Process (AHP) is a decision making model that supports complex decision making processes. This process was developed by Saaty [29], [41] and has been used extensively since. Using AHP has the advantage of allowing quantitative as well as qualitative considerations when evaluating alternative solutions [19]. AHP is powerful in that it allows the ranking of the elements and also provides the relative worth of each of the elements [19]. According to Jayaswal and Patton [19], there are multiple quality aspects that play a role in the software development process. Applying AHP using multiple quality aspects can be difficult. One option to manage the complexity is to minimise the number of the quality aspects used for evaluation of the development processes and then

to use AHP as a decision tool in this complex multi-criteria decision making process.

| e-Process Selection VBA Meta-Model Data Dictionary | | |
|---|---|---|
| **Name** | **Description** | **Type** |
| Candidate e-Processes | The e-Processes being taken into consideration. | Class |
| eProcessScores | The score for each e-Process in terms of each quality aspect as estimated by experts. | Class |
| Assessor | The expert of a specific e-Process' estimating eProcessScores. | Class |
| Quality Aspects | A list of all quality aspects including the high level quality aspects and the second level quality aspects. | Class |
| Quality Scores | The weight assigned to a specific quality aspect for a specific score. | Class |
| ProjResultsAnalysis | The result of the selection process. | Class |
| Project | Detail about the project for which an e-Process is required. | Class |
| QAAnalysis | Capturing of reasons for adjustments made to results. | Class |
| Sensitivity Analysis | The results obtained applying SA. | Class |
| Weakpoint Analysis | The results obtained applying WA. | Class |

*Table 1*. e-Process Selection VBA Meta-Model Data Dictionary

The AHP decision model is executed as follows:
- Identify the quality aspects that are going to be used to describe the problem.
- Identify the different alternative solutions to the problem.
- For each quality aspect and each pair of alternative solutions the developers give their preferences.
- Developers also need to rank the different quality aspects in terms of their significance. C1>C4>C2 etc.
- The matrices of preferences are then evaluated and a score calculated.

*1) AHP e-Process Selection Model*

The research presented in this section follows parts of the research as published in [20]. AHP is a very well-known and successfully applied method for multi-criteria decision making. Obviously the e-Process selection task is an instance of that decision making task. This method thus is justifiable. According to Ahituv and Neumann [1] Simon's model of decision making has the three stages intelligence, design, and choice. The focus here is on choice, because only existing software processes are considered. The simplified supposition is then that the decision maker identifies some existing e-Processes and the choice is made on by quantitative assessment used to identify the top-scorers and then selecting one of these.

There are a number of quality aspects which are considered essential for any method for e-Process selection. Once again these high level quality aspects and second level quality aspects are used to demonstrate the use of the analytic hierarchy process in the selection of an e-Process. The recommendation here is that the expertise required for this selection task is provided by humans in the role of either an e-Process expert or system quality expert. The AHP enables the blending of the experts' expertise to select the e-Processes best suited for the problem at hand. Choosing the experts is not necessarily trivial or cheap. The recommendation is that the qualification profiles of the designated developers are considered. If that profile indicates sufficient e-Process and system quality expertise then the recommendation is to use these developers as experts. Also recommended is that any individual with skills for both expert roles, be used – if available. Expert selection needs to consider areas like affiliation, area of competence, standing, availability, price, etc. According to Nureg-1150 [39] expert selection should consider demonstrated experience, expert versatility, expert group diversity, and expert cooperation.

*2) Quality aspect ranking and e-Process alternative ranking*

When applying AHP, there is a hierarchy structure. In this AHP e-Process selection:
- The goal is seen as the eCIS to be developed and it is placed at the root of the hierarchy structure;
- The high level quality aspects is on the second level of the hierarchy structure;
- Each of these high level quality aspect branches has associated second-level quality aspects - this is the third level of the hierarchy structure.
- The bottom level of this hierarchy structure is the set of e-Process alternatives.

Ranking occurs when e-Process experts pair-wise compare the selected quality aspects as well as rank each of the e-Process alternatives in terms of the second level quality aspects.

The quality aspects are sorted, as discussed in chapter 4, in terms of high level quality aspects and second level quality aspects. The pair-wise quality aspect comparison needs to occur within each branch of the hierarchy structure.

For the purposes of comparing two aspects x and y where x,y $\in$ { RUP, AM/XP, OSS, SBUP} or x,y $\in$ {all quality aspects} then define a predicate $\partial i$ (x,y) for i $\in$ { 1,3,5,7,9} to indicate equally preferred, moderately preferred, strongly preferred, very strongly preferred, extremely preferred, respectively.

Let H be the high level quality aspects and S be the second-level quality aspects respectively. With s $\in$ S and h $\in$ H, let Xh be the set of second-level quality aspect into which s has been decomposed with S the disjoint union of all Xh. Let q $\in$ Q = S u H.

In order to execute quality aspects ranking, the system quality experts apply the pair-wise comparison technique for assessing the relative importance of the quality aspects two times. Apply the ranking on the high level quality aspects first and then their decompositions and thus obtain weights for each.

In order to do the e-Process ranking, E is defined as the set of e-Process experts, and X a set of e-Processes. It is now

possible to denote that expert e judges $\partial i$ (x,y) = TRUE with respect to quality aspect s and define a: $S \times E \times X \times X \rightarrow \{\partial, 1/\partial \mid \partial \in \{1, 3, 5, 7, 9\}$ for a predicate t(e, $\partial$, x, y, s)and thus define: a(s,e, x, y) = $\partial$ for t(e, $\partial$, x, y, s) and a(s,e, x, y) = $1/\partial$ for t(e, $\partial$, y, x, s). This mapping is called a comparison mapping. In terms of x=y, we can define a(s, e, x, x) = 1. Similar to Kaschek [20], let be q∈Q a quality aspect, e∈E an expert, and for brevity X = {1, …, n}, then for the restriction A(q,e) of a to X × X holds A(q, e) (i, j) = a(q, e, i, j), for all i, j. This can be presented in a matrix. The elements of the matrix are the results of all pair-wise e-Process comparisons [29]:

$$A(q, e) = \begin{pmatrix} a(q,e,1,1) & ... & a(q,e,1,n) \\ ... & ... & ... \\ a(q,e,n,1) & ... & a(q,e,n,n) \end{pmatrix}$$

This matrix can be normalised and then eigenvectors for these matrixes can be determined. Solving eigenvectors consist of squaring the matrix repeatedly, each time calculating the row sums and then normalising them. Continue with this process until the difference between successive iterations is small, where small is defined beforehand. Using software designed by Raimo P. Hämäläinen [37], it is possible to employ a version of AHP, which allows for pair-wise comparisons in a consistent pair-wise comparison matrix.

*(a)      Scoring and first recommendation*

Using the mathematics from Forman and Selly [16] to calculate AHP and as defined in [20]; it is now possible to normalise and then calculate the first set of "winners". For this then use the maximum value $f^*q = \max \{fq,x \mid x \in X \}$ for normalisation, as denoted above, with the q-score wqfq,x / f*q of x as pq,x. The score px of x∈X is then defined

as $\sum_{q \in Q^2} p_{q,x} / \sum_{x \in X} \sum_{q \in Q^2} p_{q,x}$ . This then produces the set of e-Processes best suited for the problem at hand and can be defined as the set $\{x \in X \mid px = \max \{py \mid y \in X\}\}$.

*(b)      Sensitivity analysis*

Sensitivity analysis is used to identify possible poor decisions made with data input. This technique is used to identify the quality aspects and e-Processes which will, with the final ranking be the most sensitive to changes in the data. The pair-wise comparisons for both of the e-Processes and the quality aspects can be biased by the experts' subjective views and errors they might make. Sensitivity analysis is done, after the initial execution of AHP, on each e-Process quality aspect by creating a sensitivity diagram.

The idea when doing sensitivity analysis is to perform this analysis on all the quality aspects by letting one quality aspect at a time vary in small steps in an interval I. As can be seen in Figure 2, all the e-Processes are linearly represented. h2 is the diagram indicates the high level quality aspect chosen, with s1 to sn all the related second level aspects. In the diagram s2 is the selected second level quality aspect. This is the quality aspect that is depicted in the sensitivity diagram. eP1 to eP4 represents the chosen alternatives – our case possible e-Processes. Dragging the vertical line horizontally signifies

changing the quality aspect scores of the admissible e-Processes. Using this approach, then at an intersection of any of the depicted sensitivity diagrams, there can be a so-called rank reversal – another winner might be indicated.

By changing the relative importance of one of the quality aspects, there might be an impact on the overall performance of the e-Processes. By increasing/decreasing a quality aspect, there might be the impact of ending up with a different "winning" e-Process. By dragging the bar across one sensitivity diagram, there are corresponding changes to all the sensitivity diagrams [19].
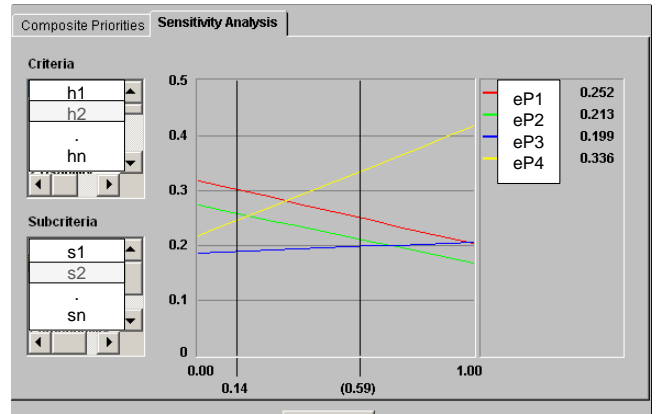


**Figure 2.** Sensitivity diagram example

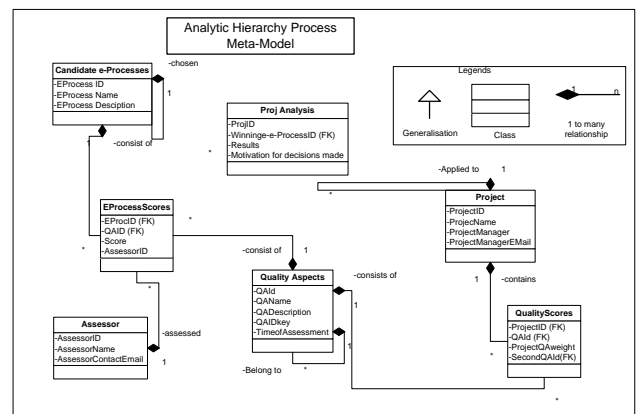*3)  E-Process Selection AHP Meta-Model*



**Figure 3.** e-Process Selection AHP Meta-Model

In Figure 3 and Table 2 the meta-model aspects for AHP are presented. When applying the Analytic Hierarchy Process then, first of all, identify the Candidate e-Processes that are going to be used to calculate the "winner" using AHP; identify and then capture those quality aspects required for the AHP calculation. Each high level quality aspect may have 0 to many second level quality aspects associated with it – to exclude a whole group of quality aspects set the top-level aspect to 0; the experts (assessors) identify the scores for each of the e-Processes and their associated quality aspects (This is then semi-permanently stored in the system) – similar to VBA; the project will for each of the quality aspects have a one to many association with the quality scores.

The quality scores will have associations with other quality scores in order to accommodate the pair-wise comparison of

AHP. After applying the AHP calculations, reasons and motivation for choices are captured in the analysis class.

| e-Process Selection AHP Meta-Model Data Dictionary | | |
|---|---|---|
| **Name** | **Description** | **Type** |
| Candidate e-Processes | The e-Processes being taken into consideration. | Class |
| eProcessScores | The score for each e-Process in terms of each quality aspect as estimated by experts. | Class |
| Assessor | The expert of a specific e-Process' estimating eProcessScores. | Class |
| Quality Aspects | A list of all quality aspects including the high level aspects and the second level quality aspects. | Class |
| Quality Scores | The weight assigned to a specific quality aspect for a specific score. | Class |
| ProjResultsAnalysis | The result of the selection process. | Class |
| Project | Detail about the project for which an e-Process is required. | Class |

*Table 2*. e-Process Selection AHP Meta-Model Data Dictionary

## C.  Case Based Reasoning Approach

Case Based Reasoning (CBR) is the approach where new problems are solved by comparing them to a number of historical cases with solutions and then choosing the "best" previous case which is nearest to this new case. The solution of the historic case then becomes, after modification, the solution of the new case. The research published in this section is based on the following publications; see [2], [3].

A number of research papers have been published on methodology selection for different domains such as frameworks for comparing object-oriented modelling (OOM) tools [21] and analysis methods [22]. They built two detailed lists of quality aspects for these selection method frameworks. Applying Case-Based Reasoning (CBR) in different domains (finance, medical, etc.) involving decision-making has been successful [38]. Access to relevant case histories of different problems reduces the requirement for problem analysis by reuse of the old problem solutions [14]. The next section discusses the CBR technique.

## 1)  Case Based Reasoning Technique Used

CBR is a computer technique, which combines development of an actual experiences knowledge base with the simulation of human reasoning. This means, search for similar situations in the past and re-use those experiences [23; 25; 42]. When a specific problem needs to be solved, similar actual occurrences are retrieved and adapted from the knowledge base. CBR is inductive (based on measuring case similarity), rather than deductive (based on logic and consistency) [14].

The four basic steps (also called Rs) of a CBR process are [14], [40], namely:

- Retrieve: When a eCIS project has been specified for which an e-Process needs to be found then retrieve a number of cases from the case database that are relevant to the problem at hand. These "old" cases will have information concerning the project that was solved, the method/s used to derive a credible e-Process and then the actual e-Process selected.
- Re-use: Once a case has been identified as being the nearest to the project at hand then adapt this old solution to fit the new project.
- Revise: Once the solution has been identified it is necessary to simulate the new solution in the real world and adapt or revise if required to fit the new project.
- Retain: Store the revised project, parameters and solution in the case database for future use.

When using CBR, the knowledge base will contain representations of a number of previous cases. A case is a contextual piece of knowledge representing a previous experience. It consists of two parts, i.e. the case-content, and the case-context. The latter tells when to use the case-content. The case-content, additionally to a case description, contains the solution applied to the problem represented by the case, and the outcome of applying the solution [40].

Figure 4 indicates the different aspects involved in the CBR cycle. Initially the cases are structured and stored in the knowledge base (also called case base). A new case that needs to be solved then gets compared to these previous cases stored in the knowledge base. Heuristic methods are used to retrieve a number of similar cases from the knowledge base. These are then compared to the new case to determine the best fit. A solution is then determined for reuse, based on one or more of these similar cases – if necessary the historical case will be revised and adapted to be a better fit. This solution will be applied to see whether it is suitable.  As soon as the solution has been validated then this solution becomes a new case in the knowledge base.
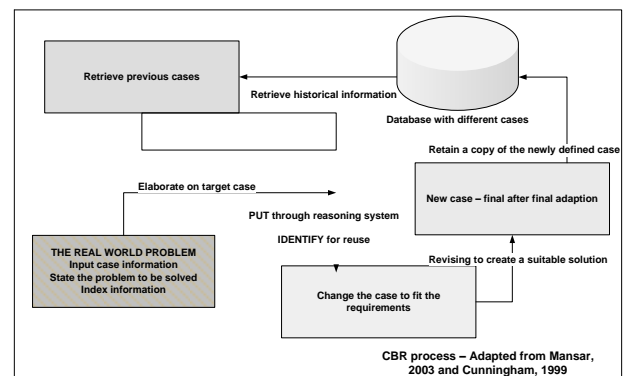


**Figure 4.** CBR process [25]

When building the knowledge base a number of methods can be used to organise, retrieve, utilise and index the past cases [36]. According to Cunningham [14] the tasks involved with building up the knowledge base are identification of the "real world" problem and representation of the key components thereof in the knowledge base. Next it is important to develop the inference mechanism that describes

the causal interactions involved in deriving solutions. The inference mechanism is implemented using the knowledge base with the cases of solved problems and a mechanism to retrieve and adapt these cases [14].

Retrieving information from the knowledge base consists of: Start with a full/partial problem description, identify a set of relevant problem descriptors, check for similarities with elements in the knowledge base, return a set of sufficiently similar cases and select the best matching case [14], [36].

Richter [27] has identified four different ways, called knowledge containers, by which knowledge can be represented in the knowledge base, namely: The vocabulary used, similarity measure, case base, and solution transformation. The CBR community has widely accepted these as a natural organisation of knowledge.

### 2) CBR E-Process Selection

It is assumed that previous cases are used in order to simplify the decision process when selecting a suitable E-Process for a specific eCIS. Case-based reasoning methods are based on the use and outcomes of past experience in order to solve a new problem [7]. A knowledge base system applies its reasoning ability through the explicit representation and use of the knowledge from a specific domain [35]. The knowledge containers are going to be discussed in the next four paragraphs in terms of case vocabulary knowledge, the use of similarity measures, case knowledge and retrieval knowledge.

CBR E-Process Vocabulary: It is recommended that when developing the knowledge base that the case vocabulary knowledge be standarised. The question that needs to be answered is whether the optimal case representation has been determined for a particular domain? It is also important to decide whether all the cases should be represented in the same way in the knowledge base. The information captured from applying the other e-Process selection decision making methods is captured in a knowledge base – see integrated meta-model.

CBR E-Process selection and the similarity measure:

A task, that needs to be executed as part of CBR, is the similarity measure. This entails that the most similar (or relevant) case/s stored in the knowledge base be identified in order to solve a particular problem, in our case to identify the most suitable e-Process to use to develop a specific eCIS. Currently our case base is still quite small and therefore all cases will be evaluated for similarity.

According to [13] the similarity between a query Q and a case C is defined as the sum of the similarities of its constituent features multiplied by the relevant weights:

$$Sim(Q,C) = \sum_{f \in F} w_f * \sigma_f (q_f, c_f).$$

In this equation $w_f$ is the constituent feature weight, $\sigma_f$ the similarity measure applied to feature f of Q and C, and F the set of all features. The weights are seen as feature attributes. The similarity measures obviously are more complex. Coyle et al. [13] use three different kinds of feature similarity measures. These are (1) the exact similarity measure, i.e., the similarity score is 1 if the feature values are equal and is 0 otherwise; (2) difference based similarity measure, i.e., the similarity score depends on the difference of the numerical feature values but

not necessarily is 0 for non-equal feature values; and (3) complex similarities, i.e., all other similarity measures.

Using the difference based similarity measure essentially turns a Case Based Reasoning problem into a version of Value-Benefit Analysis. In this method Social Choice Method (see next section) also plays a role. The author believes that the weaknesses of VBA (i.e. that it often is very hard to score items on a scale according to a number of features) can be overcome by the incorporation of SCM method parts that only rely on ranking items.

In selecting an E-Process both measure kinds (1) or (2) could be used. We are going to use (2), i.e., the difference based similarity measure. Currently the values of weights and feature similarities are obtained from the developers who were asked to assess these variables quantitatively. In future the plan is to account for the well-known critique of this approach by using SCM or AHP to obtain scores based on rankings provided by developers. In order to fully apply these methods, a software system dedicated to aid humans in applying these methods, needs to be developed.

CBR e-Process Selection Case Base: The quality of e-Processes can be seen as being multi-faceted (as is often experienced for other complex entities such as information systems [18], but individual quality aspects are often considered as too broad and unspecific and are therefore decomposed into lists of second level quality aspects. We exploit a hierarchical approach to define the eProcess by using a two level system of e-Process quality aspects – see chapter 4 (these quality aspects are used here for CBR).

Using both the VBA and AHP, a case knowledge base has been developed – see the integrated meta-model. Our Case Based has a number of levels that are used in the comparison. Firstly, there is information about each of the projects/cases being stored. Then for each of the cases we have group values (high level quality aspects) that group a number of quality aspects (second level quality aspects). There are weights assigned to each of the quality aspects.
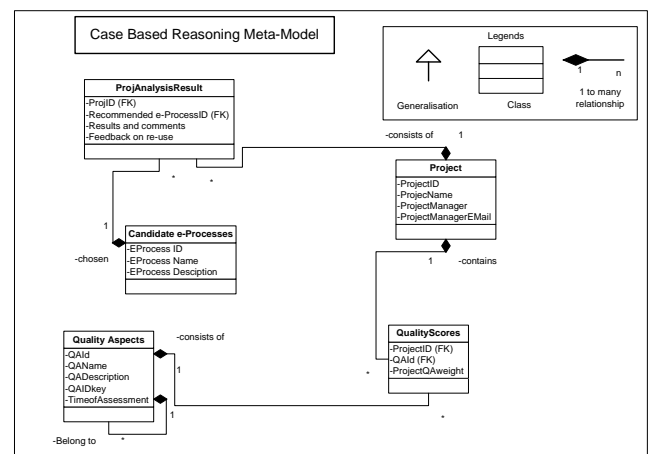


**Figure 5.** e-Process Selection CBR Meta-Model

e-Process and solution transformation: In order to determine the best match, information about the eCIS at hand is entered. This information is used in order to determine which of the cases stored in the Case Base, is the best possible solution for the problem. Once solution/solutions have been

identified, the eCIS at hand need to be transformed and becomes one of the cases being investigated.

| e-Process Selection CBR Meta-Model Data Dictionary | | |
|---|---|---|
| **Name** | **Description** | **Type** |
| Candidate e-Processes | The e-Processes being taken into consideration. | Class |
| Quality Aspects | A list of all quality aspects including the top-level aspects and the primary quality aspects. | Class |
| Quality Scores | The weight assigned to a specific Quality aspect for a specific score. | Class |
| ProjAnalysisResults | The result of the selection process. | Class |
| Project | Detail about the project for which an e-Process is required. | Class |

*Table 3*. e-Process Selection CBR Meta-Model Data Dictionary

### 3) e-Process Selection CBR Meta-Model
In Figure 5 and table 3 the meta-model aspects for CBR are presented.

### D. Social Choice Method (SCM)
The social choice method is based on voting for individual preferences from a list of candidate solutions and then aggregating the results to obtain a result that defines the collective preference [26, p. 124]. The first recorded use of a voting method was used in the voting of the Roman senate around the year 105 [17], [11]. Condorcet formulated the voting paradox and in Arrow's book of 1951 the modern form of the theory was defined [8].

There exist various social choice methods by using different acceptance criteria such as assigning decreasing points to consecutive positions and then ranking the alternatives with this point systems (Borda), Simple Majority of vote, Maximum (Score the alternatives with the worst margin they achieve and rank these according to scores), choose the ordering with minimal distance to all rankings in the profile, where the distance is defined as the number of different pair-wise relations (Kemeny method) [9]. The research in this thesis uses one of these. The detail of this method is discussed later in this section.

When applying the Social Choice Method then, first of all, identify the voters that are going to vote for the "winning" e-Process. Next, define the requirements of the new eCIS as well as those e-Processes that are going to be evaluated. Let the voters vote for their preferred e-Process. Recommend the "winner" found as the solution to which e-Process to use.

People have been using voting to decide between different choices for a long time. Social choice theory studies what different people prefer from a list of options and then aggregates the results in such a manner that a result can be obtained by defining the collective preference [26; p. 106].

The social choice method is one of the ways that voting can be applied in order to make a decision. Social choice uses the mapping of the preferences of individuals determined by an ordered list of alternatives in order to select one of the alternatives. Different methods can be used e.g. Plurality, Borda and Pair-wise comparison which is similar to AHP. In this section the focus is on using the Borda method.

According to Gaertner [17] the first recorded use of a voting method similar to the Borda method was used in the voting of the Roman senate around the year 105. The more modern format of the Borda method was by Ramon Llull (1232 – 1315) who wrote the manuscripts Ars notandi, Ars eleccionis and Alia ars eleccionis and Blanquema [17]. He is seen as the first person to document the Borda count and Condorcet criterion. The method devised by Jean-Charles de Borda in 1770 was used to elect members for the French Academy of Sciences. He published this method in Memoire sur les elections au scrutiny in the Histoire de l'Academie Royale des Sciences in Paris.

### 1) Social choice theory
According to Bernroider & Mitlohner [9] the problem of social choice can be defined in the following manner:

A set of voters n provide different rankings for the m alternatives, which results in a profile of alternatives a,b,c. If there are three voters then the rankings might be $a \succ b \succ c$, $b \succ c \succ a$, $b \succ c \succ a$. The problem requires that an aggregate raking is found $x \succeq y \succeq z$ such that the preferences of the voters are expressed in the aggregate ranking.

In social choice problems it can be assumed that the profile consists of strict orderings – but similar to Bernroider and Mitlohner [9] differences are allowed here. There are also a number of aggregate rules and one of the most important ones is the Condorcet criterion. This criterion states that if there is an alternative x that beats all other alternatives in pair-wise comparisons then this is a winner [9]. The aggregate rule should not allow a representation of weak order of alternatives or contain cycles.

Further in this voting process dimensions or attributes can be considered. The evaluation of the alternatives is now used to determine the n rakings of alternatives for the m dimensions [9].

Below find some of the major scoring procedures:
- **Simple Majority (SM):** This is a well known procedure based on margins.
  - A positive margin means that x wins against y in pair-wise comparison and results in $x \succ y$ in the aggregate relation.
  - A negative margin means that y wins against x in pair-wise comparison and results in $y \succ x$ in the aggregate relation.
  - A zero margin means indifference and results in y = x in the aggregate relation.

  This rule can result in cycles and thus limits the use of this rule in practical problems.
- **Maximin (MM):** This procedure scores the alternatives with the worst margin that they achieve and then ranks them according to these scores.

- **Copeland (CO):** This procedure scores the alternatives with the sum over the signs of the margins they achieve and ranks them according to those scores.

- **Borda (BO):** This procedure assigns decreasing points to consecutive positions.

- **Kemeny (KE):** This procedure chooses the ordering with minimal distances to all rankings in the profile, where the distance is defined as the number of different pair-wise relations.                     [9]

The plan is to use one of the Borda procedures therefore this section continues to discuss some of the Borda aspects. Borda, one of the social choice methods, initially starts by ranking each alternative solution for a given criteria by giving the preferred alternative a value of 1, the second alternative a value of 2, the third one a value of 3 and so on. The social choice, or aggregate pre-order, is calculated by calculating the total value for each of the alternative solutions. This approach means that the winner is the solution with the least points.

According to Pomerol and Barba-Romero [26, p.106] m integer alternatives are chosen such that $k1 > k2 > k3 > ... km \geq 0$. These are called the Borda coefficients. For each of the different criterion j the alternatives are ranked according to a complete pre-order. These rankings are called rij (where alternative i is associated with the pre-order associated with criterion j. Borda voting, according to Pomerol & Barba-Romero [26; p.106], is thus for a given n complete preorders $\succsim$j (denotes strict preference or indifference) on m alternatives A1, A2, …, Am it is the procedure which, for a given alternative Ai , consists of taking the sum of the votes (or $\sum k \neq I$ vik) that it obtains in all possible duels of Ai versus Ak. The alternatives are then ranked in order of the number of votes.

When using Borda all the personal preferences are weighed and the highest scoring one is the winner. Even though Borda can be seen as the peoples' choice the scoring method uses cardinal utilities and disregards personal preferences. Borda focuses on the complete reference profile [28], [44]. One problem that can be identified with majority voting is that only the top preferences are taken into account [28], [44]. Another method that can be taken into account is the adjusted Borda method [34]. This method basically consists of counting the votes for and the votes against a specific candidate and calculating the difference.

*2)   The SCM e-Process Selection Methodology*
As explained above, when applying the Borda rule, decreasing points are assigned to consecutive positions, such that in our case 3 points is assigned to the first place, 2 points to the second place, 1 point to the third place and zero to the fourth. In this instance the aim is to choose between the e-Process quality aspects in order to select an e-Processes.

*3)   2.4.3  e-Process Selection SCM Meta-Model*
The developers will be asked to vote on each of the high level aspects used for evaluating the environment as well as which of the e-Processes they prefer to use for the development

process. This information is then evaluated in order to determine which one will be best suited. In Figure 6 and Table 4 the meta-model aspects for Social Choice method is presented.
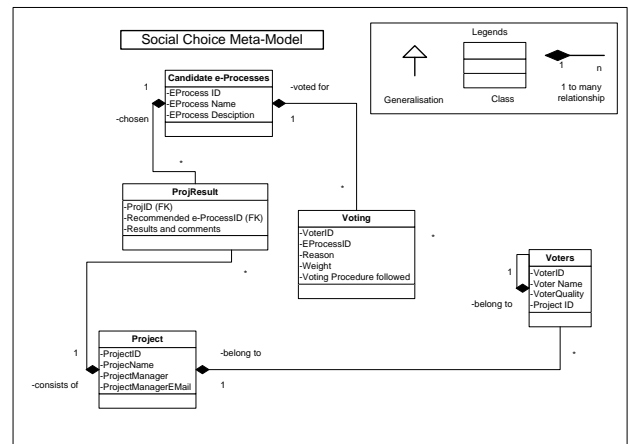


**Figure 6.** e-Process Selection SCM Meta-Model

| e-Process Selection SCM Meta-Model Data Dictionary | | |
|---|---|---|
| **Name** | **Description** | **Type** |
| Candidate e-Processes | The e-Processes being taken into consideration. | Class |
| Voting | A list of all votes obtained. | Class |
| Voters | Detail about all voters participating in the process. | Class |
| ProjAnalysisResults | The result of the selection process. | Class |
| Project | Detail about the project for which an e-Process is required. | Class |

*Table 4:* e-Process Selection SCM Meta-Model Data Dictionary

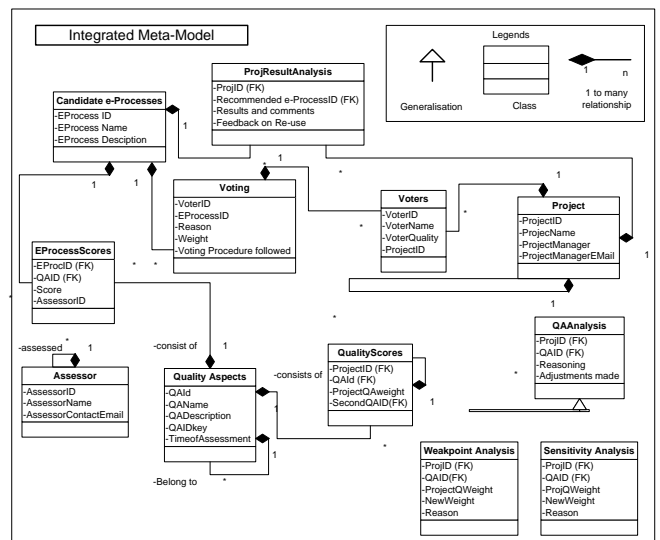## III.  Integrated Meta-Model



**Figure 7.** Integrated e-Process Selection Meta-Model

In Figure 7 and Table 5, the meta-model aspects, for a combined view of all four decision models are presented.

When applying any of these decision methods the resulting meta-model should be able to accommodate all of these. This journal paper introduced some of the basic concepts required in the research. It introduced some of the e-Processes that will be used in the models, along with some of the decision models that will form the basis of the methodology used in this study.

| Integrated e-Process Selection Meta-Model Data Dictionary | | |
|---|---|---|
| **Name** | **Description** | **Type** |
| Candidate e-Processes | The e-Processes being taken into consideration. | Class |
| eProcessScores | The score for each e-Process in terms of each quality aspect as estimated by experts. | Class |
| Assessor | The expert of a specific e-Process' estimating eProcessScores. | Class |
| Quality Aspects | A list of all quality aspects including the high level aspects and the second level quality aspects. | Class |
| Quality Scores | The weight assigned to a specific quality aspect for a specific score. | Class |
| ProjResultsAnalysis | The result of the selection process. | Class |
| Project | Detail about the project for which an e-Process is required. | Class |
| QAAnalysis | Capturing of reasons for adjustments made to results. | Class |
| Sensitivity Analysis | The results obtained applying WA. | Class |
| Weakpoint Analysis | The results obtained applying WA. | Class |
| Voting | A list of all votes obtained. | Class |

*Table 5:* Integrated e-Process Selection Meta-Model Data dictionary

## References

[1]   N. Ahituv & S. Neumann. "Principles of information systems for management" (3rd ed.) Wm. C. Brown Publishers. 1990..

[2]   F. Albertyn & R. Kaschek. "A hybrid method for E-Process selection". In proceedings of *18th Australasian Conference on Information Systems (ACIS 2007)*. Toowoomba, Australia. (http://www.acis2007.usq.edu.au/). 2007.

[3]   F. Albertyn. "E-Process Selection using Case-Based Reasoning Techniques". In proceedings *of IADIS international conference E-Society 2007*. Lisbon, Portugal. 2007. ( http://www.esociety-conf.org/) part of *the IADIS Multi Conference on Computer Science and Information Systems (MCCSIS 2007)* Lisbon, Portugal, 3 - 8 July 2007 (http://www.mccsis.org) .

[4]   F. Albertyn. "Ontology for the Selection of e-Processes". In *proceedings of Web Information Systems Engineering - WISE 2005 Workshops*. pp. 1 - 10. New York, eds. G. Y. Dean M., Jun W., Kaschek R., Krishnawasmy S., Pan Z., Sheng M. 2005.

[5]   F. Albertyn, R. Kaschek. " eCIS Development: Refining E-Process Selection". In proceedings *of the IADIS International Conference - e-Society*, Qawra, Malta, P. K. a. M. McPherson (eds), pp. 233 – 240, ISBN 972-8939-03-5. 2005.

[6]   F. Albertyn, R. Kaschek, R. "Ontology Based E-Process Selection". In the proceedings *of the Sixth International Conference on Information Integration and Web-based Applications & Services (iiWAS2004)*, Jakarta, Indonesia, Austrian Computer Society, pp. 143 – 152. G. Kotsis, S.Bressan, D. Taniar and I. K. Ibrahim eds.. ISBN. 3-85403-183-01. 2004.

[7]   E. Armengol, S. Onta˜non, & E. Plaza. "Explaining Similarity in CBR". Paper presented at *the ECCBR'04*, Madrid, Spain. 2004.

[8]   K.J. Arrow. "Social Choice and Individual Values". ISBN 0-300-01364-7. 1963.

[9]   E.W. Bernroider, J. Mitlohner. "Resolving Indifferences in Multiple Attribute Business Information System Selection with Social Choice Methods". In the proceedings *of the 6th International Conference ISTA 2007*, Kharkiv, Ukraine. 2007.

[10]  R. Böhm, S. Wenger. "Methoden und Techniken der System-Entwicklung". Zürich, vdf Hochschulverslag AG and der ETH. 2nd. ed. 1996.

[11]  Borda-Wikipedia. "Borda Count". Retrieved in March 2008 from http:/en.wikipedia.org/wiki/Borda_count. 2008.

[12]  P.L. Clemens. "Weighted Scoring Decision Making". Jacobs Sverdrup. 2002.

[13]  L. Coyle, D. Doyle, et al. "Representing similarity for CBR in XML". In proceedings of the *Advances in case-based reasoning ECCBR 2004 : European conference on case-based reasoning No7*, Madrid, ESPAGNEISBN 3-540-22882-9. 2004.

[14]  P. Cunningham, A. Bonzano. "Knowledge engineering in a real world casebased reasoning application". Technical Report TCD-CS-1999-36, Trinity College, Dublin. 1999.

[15]  M. Doumpos, C. Zuopounidis. "Multicriteria Decision Aid Classification Methods". Dordrecht: Kluwer Academic Publishers. 1-4020-0805-8. 2002.

[16]  E.H. Forman, M.A. Selly. "Decision By Objectives: How to convince others you are right". World Scientific Press. Retrieved on 3 March 2008 from http://mdm.gwu.edu/forman/DBO.pdf. 2001.

[17]  W. Gaertner. "Domain conditions in social choice theory". Cambridge, UK: Press sindicate of the University of Cambridge. 0-521-79102-2. 2002.

[18]  C. Ghezzi, M. Jazyeri, et al. "Software qualities and principles". In: Tucker, A.B.: Computer Science Handbook. Chapman & Hall/CRC. 2004.

[19]  B.K. Jayaswal, P.C. Patton. "Design for trustworthy software. Tools, techniques and methodology of developing robust software". United States: Prentice Hall Pearson education. 2006.

[20] R. Kaschek, F. Albertyn, V. A. Shekhovtsov, S. Zlatkin. "An e-Process Selection Model". In the Proceedings of *the Twenty-Sixth International Conference on Conceptual Modeling - ER 2007 - Tutorials, Posters, Panels and Industrial Contributions*, Auckland, New Zealand. 5 - 7 November 2007. Conferences in Research and Practice in Information Technology, Vol. 83. John Grundy, Sven Hartmann, Alberto H. F. Laender, Leszek Maciaszek and John F. Roddick, Eds. 2007.

[21] R. Kaschek, H.C. Mayr. "A Characterization of OOA Tools". *Proceedings of the Fourth International Symposium on Assessment of Software Tools (SAST '96).* 1996.

[22] R. Kaschek, H.C. Mayr. "Characteristics of O-O Modeling Methods". In the proceedings of *the EMISA Forum*, 1998(2), pp. 10-39. 1998.

[23] D.B. Leake. "CBR in context: the present and future. Case Based reasoning experiences - Lessons and future experiences". D. Leake. Cambridge, MIT Press. 1996.

[24] S.J. Mantel Jr., J.R. Meredith, S.M. Shafer, M.M. Sutton. "Project. Management in Practice". Hoboken: John Wiley & Sons. 2005.

[25] S.L. Mansar, F. Marir, H. Reijers. "Case-Based Reasoning as a Technique for Knowledge Management in Business Process Redesign". Academic Conferences Limited. Retrieved in June 2003 from www.ejkm.com. 2003.

[26] J.C. Pomerol, S. Barba-Romero. "Multicriterion Decision in Management: Principles and Practice". MassaChusetts: Kluwer Academic Publishers. 0-7923-7756-7. 2003.

[27] M.M. Richter. "Introduction - the basic concepts of CBR. Case-based reasoning technology: From foundations to applications". M. Bartsch-Sporl and S. Wess, Springer-Verlag. LNCS vol. 1400. 1998.

[28] A. Riedl. "Voting and Social Choice" Chapter 4. Retrieved in April 2004 from www1.fee.uva.nl/creed/pdfiles/moulinch4elferink.pdf. 2004.

[29] T.L. Saaty. "Multicriteria Decision Making: The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation". Pittsburgh: RWS Publications. 1990.

[30] R.E. Trueman, R. E. "An Introduction to quantitive methods for decision making". (Second edition ed.): Library of Congress Cataloging in Publication Data. 0-03-018391-x. 1976.

[31] P. Fettke, P. Loos. "Ontological evaluation of reference models using the Bunge-Wand-Weber model". Ninth Americas Conference on Information Systems. 2003.

[32] A.L. Opdahl, B. Henderson-Sellers, F. Barbier. "Ontological analysis of whole-part relationships in OO-models". *Information and Software Technology*. 43(2001). pp. 387 - 399. 2001.

[33] Y. Wand, V.C. Storey, R. Weber. "An ontological analysis of the relationship contruct in conceptual modeling". *ACM Transactions on Database Systems*. 24(4). pp. 494 - 528. 1999.

[34] D. Black. "The theory of committees and elections". Boston: Kluwer. 1958.

[35] B. Diaz-Agudo, P.A. Gonzalez-Calero. "An architecture for Knowledge Intensive CBR systems". Retrieved in April 2005 from www.idi.ntnu.no/~agnar/ tdt55/ 2000_ewcbr_belen-2.pdf. 2000.

[36] I. Harrison. "Case Based Reasoning. Inductive Solutions". Retrieved in August 2006 from http://www.aiai.ed.ac.uk/links/cbr.html. 1997.

[37] R.P. Hämäläinen. "Web Hibre – Global Decision Support Software". Copyright 1998-2007 Systems Analysis Laboratory, Helsinki University of Technology. All rights reserved. Accessed on web in October 2008 at http://www.hipre.hut.fi/. 2008.

[38] D.J. Macaya, J. Melendez, et al. "Case Based approach for supervision. Application to PID controllers". In the Proceedings of the *15th-IFAC World Congress*, Barcelona. 2002.

[39] NUREG-1150. "Reactor Risk Reference Document". U.S. Nuclear Regulatory Commission, Washington D.C. 1989.

[40] L. Watson, F. Marir. "Case-Based Reasoning: A Review". Retrieved in July 2006 from www.surveying.salford.ac.uk/ai-cbr-mirror/classroom/cbr-review.html. 1994.

[41] T.L. Saaty. "Principia Mathematica Decernendi: Mathematical Principles of Decision Making". Pittsburgh, Pennsylvania: RWS Publications. ISBN 978-1-888603-10-1. 2010.

[42] J. Prentzas, I Hatzilygeroudis. "Combinations of case-based reasoning with other intelligent methods". *Int. J. Hybrid Intell. Syst*. 6, 4 (December 2009). pp. 189-209. 2009.

[43] M. Goldmanis, A. Hortaçsu, C. Syverson, Ö. Emre. "E-Commerce and the Market Structure of Retail Industries". *The Economic Journal*, Volume 120, Issue 545, pp. 651–682. June 2010.

[44] D. Bouyssou, D. Dubois, M. Pirlot, H. Prade, T. Marchant, P. Perny. "Chapter 19. Social Choice Theory and Multicriteria Decision Aiding. Decision-making Process" Concepts and Methods. Published Online: 19 JAN 2010.Wiley Online Library. 2010.

[45] C. Yang, M. Wu, C. Lin, D. Yang. "The Implementation of Wiki-based Knowledge Management Systems for Small Research Groups". *International Journal of Computer Information Systems and Industrial Management Applications (IJCISIM)*. ISSN 2150-7988 Vol.1 (2009), pp. 68–75. http://www.mirlabs.org /ijcisim. 2009.

## Author Biography

**Frina Albertyn** was born in Paarl, South Africa. She graduated with a BSc in Mathematics and Computer Science in 1979 from the University of Pretoria, followed by a BSc(Hons) in Computer Science from the University of South Africa. She completed an MSc in Computer Science at the Rand Afrikaans University in 1997. Frina immigrated to New Zealand in 1999. She graduated with a PhD in Information Systems in 2010 from Massey University in Palmerston North. She has published in several conference proceedings and journal papers internationally and nationally on a number of topics**.**