QUALITY MEASURES AND THE INFORMATION CONSUMER

(Research in Progress)

M S E Burgess, W A Gray, and N J Fiddian
School of Computer Science, Cardiff University, Wales, UK.
{m.burgess, w.a.gray, n.j.fiddian}@cs.cardiff.ac.uk

Abstract: In recent years the amount of data available to the information consumer has dramatically increased. It is now possible to search for information on an unlimited number of topics across a wide range of information environments. Although plentiful, this information is also of varying levels of quality, being produced both by professionals and those with little or no subject knowledge. As such, it is becoming increasingly difficult to find precisely what is required. The two hurdles that prevent the finding of relevant information are therefore 'information overload' and 'information quality'.

Our proposed solution to this problem consists of the development of a methodology for using quality criteria as an aid to information searching. Having previously developed and presented a generic hierarchical framework of quality, and corresponding domain-specific frameworks, we now demonstrate how these models can be used by the information consumer.

Using our experimental Information Search Environment the information consumer is able to create a personalised definition of quality, based on the selection of quality criteria, importance weightings, and quality level preference values. This quality definition is then used to focus information searches in their chosen subject domain. In this paper we present our approach, and show how changing this quality definition can alter the results returned from an information search.

INTRODUCTION

In recent years the volume of data readily available to the information consumer has dramatically increased. It is now possible to search for information on an unlimited number of topics across a wide range of information environments, such as electronic library systems, corporate intranets and the Internet. Although plentiful, this information is also of varying levels of quality, with providers ranging from multi-national corporations to individuals with limited knowledge. This range of suppliers also results in a diverse variety of formats in which information is stored and presented.

With so much information available, quality has become an important discriminator when deciding which information to use and which to discard. However, problems such as information overload, specificity of database queries, and the requirement for users to be able to explicitly state their information need, can hinder their search for information that meets their current individual need.

Due to the large amount of data now available to the information consumer, and being comparatively easy to access, an assumption could be made that finding information on a desired topic should be a straightforward task. However, due to the amount of information being so large, and being of varying levels of quality, it is becoming increasingly difficult to find precisely what is required, particularly if the information consumer does not have precise knowledge of their information needs. The two hurdles that prevent the finding of relevant information are therefore 'information overload' and 'information quality'. Our proposed solution to this problem consists of the development of a methodology for using quality criteria as an aid to information searching.

Research Premise

The hypothesis of our research asserts that it is possible to create a hierarchical generic model of quality that can be used by the information consumer to assist in information searching, by focusing the returned result set based on personal quality preferences.

The first part of this paper discusses definitions of quality, including our hierarchical generic quality framework and the domain-specific quality frameworks derived from the generic model. It then continues by discussing the experimental Information Search Environment created to demonstrate how the information consumer can use quality to focus information searches across a number of domains, and how changes in quality preferences produce different result-set ranking orders.

DEFINING QUALITY

Although people intuitively know what is meant by the term 'quality', when asked to produce an explicit definition most will struggle. This is the principal problem that is encountered when discussing quality: everyone knows what it is but very few people can define it. This leaves us in a difficult position when wanting to incorporate quality in some computational system, as to be used in this type of environment an explicit definition with quantitative representations of terms is essential. Therefore, before quality can be incorporated into such a system an explicit definition must be obtained.

Although quality is a difficult term to define, some research has been conducted in this area. Table 1 demonstrates the variety of quality definitions and models that have been developed, across a selection of example subject domains: Software Quality, Data Quality, Information Quality, and Web Quality. As can be seen in this table, although research has been conducted into creating a definition of quality "*no single definition or standard of quality exists.*"[36].

PROJECT	YEAR	DOMAIN	QUALITY FRAMEWORK STRUCTURE
SOFTWARE QUALITY DEFINITIONS			
Barbacci et al [3]	1995	Software quality	4 models for each of 4 primary attributes, with a total of 13 concerns
Boehm et al [4]	1976	Software quality	Hierarchical tree structure comprising 10 categories and 15 metrics
Dromey [13;14]	1995	Software quality	3 models, containing 17 attributes and 42 unique sub-attributes (repeated amongst the models)
Hyatt & Rosenberg [21]	1996	Software quality	4 goals and 13 attributes
ISO 9126-1:2001 [22]	2001	Software quality	2 models: 1) 'Internal & external software qualities' - 6 dimensions & 34 metrics. 2) 'Quality in use' – 4 metrics
Liu et al [23]	2000	OO software design	3 factors and 8 criteria
McCall [26]	1977	Software quality	3 classes, 11 factors, and 23 criteria
Ortega et al [31]	2001	Software quality	6 metrics
Royce [34]	1990	Software products	4 metrics
Ruby & Hardwick [35]	1968	Software quality	7 attribute descriptions
DATA QUALITY DEFINITIONS			
Abate et al [1]	1998	Data quality	4 categories and 15 dimensions
Cykana et al [11]	1996	Data quality	6 characteristics
Gardyn [18]	1997	Data warehouse	5 dimensions
Long & Seko [24]	2002	Medical data quality	5 dimensions and 24 characteristics
Naumann [28]	2002	Query planning	4 dimensions and 22 metrics

Redman [33]	1996	Data quality	3 categories and 27 dimensions
Wang & Strong [38]	1996	Data quality	4 categories and 15 dimensions
INFORMATION QUALITY DEFINITIONS			
Bovee et al [5]	2001	Information quality	4 criteria and 10 components
Dedeke [12]	2000	Information systems	5 dimensions and 28 metrics
Eppler [15]	2001	Information quality	4 quality levels and 16 criteria
Matsumura & Shouraboura [25]	1996	Information quality	2 categories and 4 attributes
Miller [27]	1996	Information quality	10 dimensions
WEB QUALITY DEFINITIONS			
Aladwani & Palvia [2]	2002	Site quality	4 dimensions and 25 items
Chen et al [9]	1998	Query processing	10 quality parameters
Olsina et al [30]	2001	Academic sites	Hierarchical model containing 100+ metrics
Zhu & Gauch [41]	2000	Site quality	15 metrics

Table 1 Selection of current definitions of quality

Although Table 1 shows a variety of definitions, using different terminology (such as 'attributes', 'criteria' and 'metrics'), all identify:

- the importance of a definition of quality; and
- that quality is a multi-attribute entity.

The majority also agree that the multiple attributes used to define quality can be grouped into related categories, representing a hierarchical structure.

The Information Consumer Perspective

Most of the work currently conducted in the area of quality research has looked at quality from the organisational or information producer perspective. The information consumer's perspective of quality differs from these in two important ways:

- The consumer has no control over the quality of available information.
- The aim of the consumer is to find information that matches their personal needs, rather than provide information that meets the needs of others.

This difference in focus means quality definitions that have been defined for use by information providers are not suited to the information consumer.

The typical information consumer wants to find the best available information that meets their requirements, at that point in time, in their current domain of interest. This may not necessarily be the best possible result as the consumer often has restrictions, such as the time available to spend searching for information. For example, the consumer may need the information quickly so is unable to wait several hours while all possible sources of information are investigated to find the best result across all sources. In this case the consumer will be willing to accept the best possible results obtainable within the given restriction, such as currently available data, data within their price range, or all data that can be obtained within a specified time limit.

A Consumer-Oriented Definition of Quality

In earlier papers we presented a hierarchical generic model of information quality. Based on previous research into defining quality, consumer perceptions of quality, and definitions used by consumer organisations, this model was developed to assist the information consumer in creating a personalised definition of quality. Below we summarise our quality framework, but for more information on how it was developed the reader is referred to [7] and [8].

A Hierarchical Generic Framework of Quality

The hierarchical generic framework of quality developed during the initial phase of this project can be seen in [7]. The criteria identified for incorporation in this model were based primarily on those identified during an investigation into previous definitions of quality, and features of quality stated as important by consumer representation organisations.

The principal features of this framework of quality are as follows:

- Generic – contains a set of quality criteria applicable to a range of subject domains.
- Hierarchical – allows criteria to be grouped into related categories, and sub-categories.
- Intuitive – criteria needs are locatable by following an intuitive path through the hierarchy, to ensure its ease of use.
- Flexible – (i) provides facilities for the relocation of criteria within the hierarchy, (ii) enables the selection of a set of criteria to create a personalised definition of quality.
- Extensible – provides facilities for the addition of new criteria and the deletion of un-required criteria.

Domain-Specific Frameworks of Quality

To be able to demonstrate how quality attributes can be used to assist in information retrieval the generic framework needs to be focussed onto a real world domain. This requirement results in the need for domain-specific quality criteria. Although these criteria could be incorporated into the initial generic framework, after the inclusion of just two or three different domains the number of criteria would become unmanageable.

Our proposed solution to this is the use of the generic framework as a blueprint for the development of separate, domain-specific frameworks. By using the structure of the generic framework as a starting point, relevant generic criteria can be selected followed by the inclusion of criteria specific to the chosen domain. This domain-specific framework can then be used to facilitate a search for information within that topic.

To facilitate the creation and maintenance of both the generic and domain-specific quality frameworks we created the Quality Toolkit – a proof-of-concept application to demonstrate framework creation, maintenance, flexibility and extensibility. An example of a domain-specific framework based on our original generic model can be seen in [8], in which we present a quality framework for the UK university domain.

USING QUALITY TO FOCUS INFORMATION SEARCH RESULTS

When searching for information, especially in multiple provider environments, the user can become inundated due to the vast number of potential results. This is particularly the case when conducting a search on the Internet. As the number of information items available increases the consumer starts to suffer from information overload, where they are no longer able to effectively process that information, as illustrated in Figure 1.

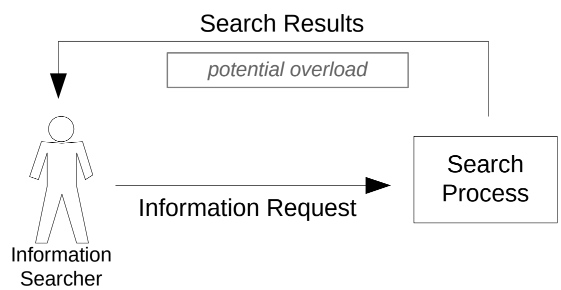

Figure 1 Potential for information overload in an information search

We propose that by developing a model of quality that can be used to create a personalised definition of quality, according to the individual consumer, this quality profile can be used to focus an information search onto a relevant set of results, thus reducing information overload and increasing consumer productivity, as in Figure 2.

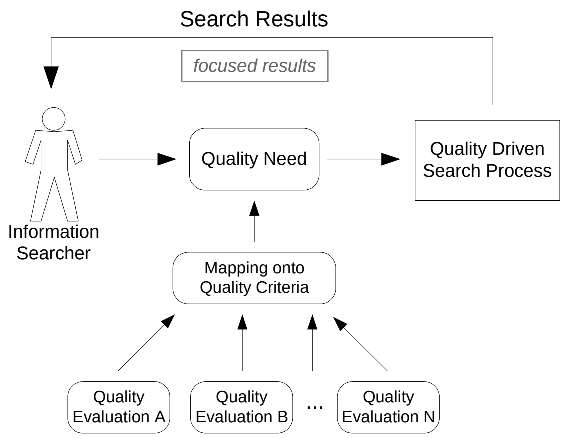

Figure 2 Information search focused according to quality requirements

The focus of our previous paper [8] was the creation of a framework of quality that can be used to assist the user in stating their individual quality need, which can then be used in a quality-driven search process. It also discussed the currently manual approach to mapping available data to the quality criteria (discussed below).

In this paper we concentrate on the exploitation of the resulting quality frameworks, discussing how quality criteria evaluations are obtained, and how quality can be used to focus an information search based on user preferences.

QUALITY CRITERIA EVALUATIONS

Before quality can be used as an aid to information searching, methods need to be defined for obtaining values for the identified quality criteria.

Classes of Quality Criteria Assessment

In his 2000 paper, Naumann [29] defines a set of classes for the assessment of information quality (IQ), and identifies methods developed for use in each class. The sources for quality criterion evaluation identified in his paper are the **user**, the **information source**, and the **query process** used when obtaining the information. These three sources are divided into the following classes of assessment:

Subject criteria - when IQ scores can only be obtained from individual users, based on their personal views, experiences, and background.
Object criteria - when IQ scores can be obtained by analysing the information.
Process criteria - when IQ scores are determined by the query process.

For each of these assessment classes Naumann presents a set of assessment methods that can be used to evaluate the quality of each information source:

Subject criteria - user experience, user sampling, continuous user assessment.
Object criteria - contract of content quality, content parsing, content sampling, expert input, continuous assessment of content.
Process criteria - data cleansing, continuous assessment of process, structural parsing.

For a comprehensive explanation of these methods of assessment the reader is referred to Naumann's book "Quality-Driven Query Answering for Integrated Information Systems" [28].

Source vs. Information Quality

During Naumann's work the emphasis was on obtaining values for the quality of the information source, and the process used for accessing those sources, such as query processing quality. The thrust of our work was to ascertain the quality of the available information, rather than the source from which the information is obtained. Our work is therefore complementary to the work presented by Naumann and his colleagues: as whereas Naumann's approach looks at the coarse-grained aspect of source quality, ours focuses on the fine-grained aspect of the individual information items.

Due to the differing focus of our two approaches the assessment methods described above are not immediately transferable to our research. To be of use in our work they must be modified, to consider quality of the information rather than the source from which it is obtained. However, this modification is only to the implementation of the assessment methods; their definitions remain the same.

Frequency of IQ Criteria Assessment

An important issue when automatically obtaining IQ scores, and storing them for use again at a later time, is the frequency at which these values require updating. This is dependant on whether a criterion is static or dynamic in nature.

If a criterion is static then once a value is obtained and stored it will be possible to use that same value for some considerable period of time, and only check for changes after an appropriate time lapse. However, for dynamic criteria a judgement must be made as to how frequently these updates need to occur. If a criterion is dynamic over a period of weeks, then its value only needs to be updated after a predetermined number of weeks. If, on the other hand, the value can change in hours, minutes, or seconds such as when considering values of stocks and shares a frequent update cycle will be required, which has an elapsed time related to the time between expected changes.

The quality criteria used in our experimental research domains have been considered as static, in that once evaluated the values for these criteria do not change. Although in a real-world system a number of criteria would be dynamic, with evaluations requiring frequent updating, this dynamic nature was removed to ensure consistency of experimental results. The removal of dynamic criteria values does not adversely affect the results presented in this paper as we are only currently concerned with static criteria, and dynamic values will be the focus of further work.

Experimental Subject Domains

To demonstrate how quality can be used in the search for information within some data set, the generic quality framework was used as a basis for the creation of a set of domain-specific frameworks. Developing these sub-frameworks makes it possible to demonstrate the use of quality in real-world domains, with data obtained from real-world sources.

Two primary subject domains were selected for experimentation:

- UK Universities
- Cars

The reason for the selection of these domains is to show how domain-specific quality frameworks can be developed and used when searching for a service (e.g. universities), and information on a tangible product (e.g. cars). Other subject domains have also been implemented for a number of tangible products including home freezers, luggage and cameras.

All of the data for each of these subject domains were acquired from Internet sources. The majority of these sources provide information via structured web sites, so although the data was extracted manually for use in our experiments, it is also feasible to employ automatic parsing techniques to extract this data.

Mapping Quality Criteria to Available Data

To be able to calculate quality scores for available information items we need to know from where values are obtained for each quality criterion. This requires a mapping between quality criteria and the available data (i.e. values obtained from external sources which can be used to calculate quality scores for each item or piece of information). This mapping is currently created manually, using the Quality Toolkit – a tool created during this project to create and maintain quality frameworks. Using this tool each data field is mapped to the appropriate quality criterion, along with a ranking of its importance when calculating the criterion value. This is necessary due to each criterion score typically being based on a number of values, from more than one source. For example, in the university domain the criterion of ‘Teaching Quality’ is based on teaching data values obtained from both of the employed data sources (The Times [37] and The Guardian [19]).

As this is currently a manual process the potential exists for further research into developing a semi-automatic process for creating criterion mappings. However, the method for mapping criteria to available data does not affect how the information consumer uses quality to focus information searches.

QUALITY-DRIVEN INFORMATION SEARCHING

To demonstrate how quality criteria can be used to assist in information searching, the experimental Information Search Environment (ISE) was developed. This proof-of-concept application incorporates the ideas presented in this paper, whereby quality criteria can be selected from within a chosen quality framework and then used to focus information search results.

Quality Preferences

A search for information using ISE comprises five stages:

1. Selection of domain of interest
2. Selection of set of quality criteria to focus the search
3. Stating importance weightings for each selected criterion
4. Stating preference values, if desired, for each criterion
5. Stating importance rankings, if desired, for information providers

In this section we discuss each of these stages, and then conclude with a discussion on how results can be ranked based on user-specified quality requirements.

Selection of Subject Domain and Quality Criteria

A number of subject domains are currently available for selection by the user, in which to conduct an information search. The principal domains we have chosen for experimentation, as discussed above, are those of UK universities and new cars.

Once the user has selected their domain of interest they are presented with the corresponding domain-specific quality framework, from which they can select a number of quality criteria. Figure 3 shows the criteria selection process when the university domain has been chosen.

As well as selecting those criteria the user feels are important the user can also opt to select a default set of quality criteria. This feature enables popular quality criteria to be automatically selected for use in the current search, with the option for the user to remove any of these criteria and include others. In our demonstration environment default criteria are identified manually using the Quality Toolkit. However, this has been identified as/an area requiring further research to enable these default criteria selections to be generated automatically based on previous user selections and explicit user feedback.

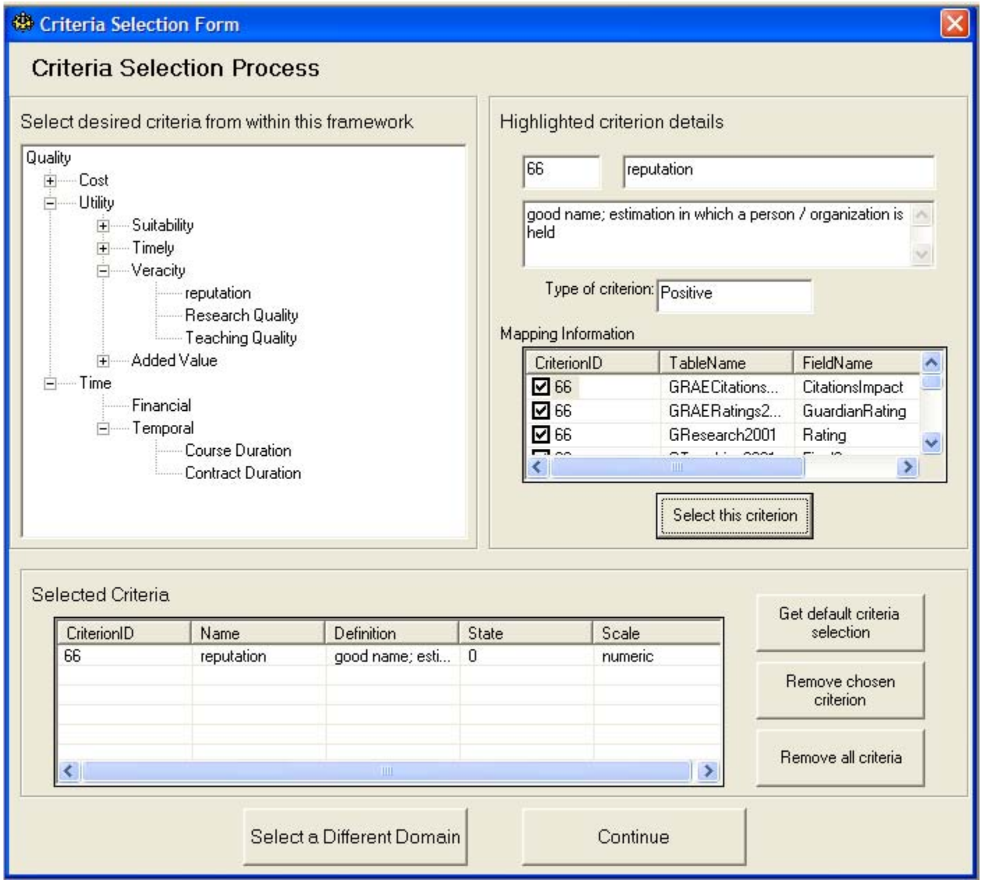

Figure 3 Selecting quality criteria in the university domain.

Importance Weightings

Once a set of quality criteria have been selected, the user is required to state weighting values for each criterion. These importance weightings are obtained from the user by presenting them with a graphical slider for each of their chosen criteria, and asking them to rate criterion importance using a percentage scale. They are then used by the ranking algorithm (see below) to find the best possible results, based on the importance of each criterion to the user, in their current situation.

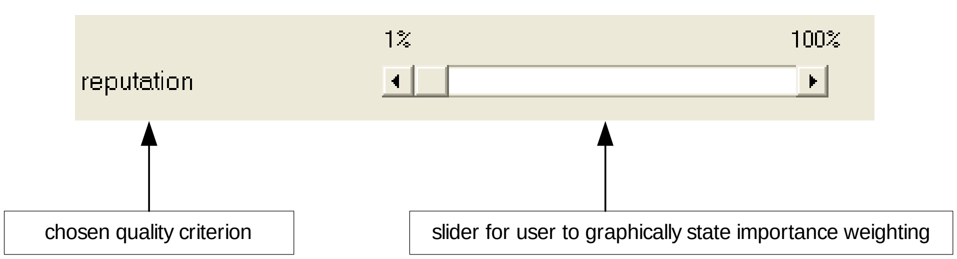

Figure 4 Example graphical slider.

Preference Values

As well as stating importance weightings the facility also exists for the user to state preferred values for each criterion. Although the highest value for a quality criterion might be assumed to be the most preferable, this is not always the case. The user may not wish to search for the items that have the highest values for their chosen criteria. For example, they may wish to trade-off against cost and therefore search for some item valued up to a maximum price. These preference values are obtained from the user in the same way as importance weightings: via graphical sliders for each criterion.

User-Stated Source Preferences

Data for the two primary experimental subject domains were obtained from multiple sources:

- UK universities: The Times [37] and The Guardian [19] (two UK national newspapers).
- New cars: Euro NCAP [17], Parkers [32], What Car [39], and Which [40] (independent consumer organisations).

When combining information from a variety of different sources, or multiple databases within a single environment, many difficulties arise. These include:

- synonyms and homonyms,
- multiple formats of data representation and storage,
- data repetition,
- incorrect data,
- conflicting data.

While obviously a problem when dealing with data in the real world, this is being investigated by other researchers and is a research area in its own right. Although the data used in our project has been obtained from multiple sources, it is stored in a local database and has gone through manual data checking to resolve the aforementioned difficulties.

This local storing of data also ensures consistency between experimental results. If the data were accessed directly from their original source we would have no control over that data. The information providers could potentially change their data, thus making it difficult to draw conclusions from our experiments into quality-oriented searching.

The current version of ISE allows the user to state data source preferences when quality criteria evaluations come from multiple sources. For example, when searching for information on UK universities the user can state whether information obtained from ‘The Times’ is preferable over information from ‘The Guardian’ by setting the ranking orders to *first* and *second* respectively. It is also possible to eliminate undesired data sources by setting their ranking value to 0. These ranks are then used as a weighting value for data obtained from each source when calculating values for each quality criterion.

Ranking of Search Results

To find the best results based on user stated criteria settings, from the available data, three ranking algorithms have been employed: SAW, TOPSIS and TOPSIS-GP.

Simple Additive Weighting (SAW)

SAW is the best-known and most widely used multi-attribute decision-making (MADM) ranking method [20]. This algorithm comprises three basic steps:

Step 1: Scale quality criteria values v and weights w using transformation functions.

Step 2: Apply the scaled user defined weighting to each quality criteria

Step 3: Calculate quality score iq for each item S_i by summing scores for each criterion.

The final score for each data item is therefore calculated as follows:

$$iq(S_i) \Rightarrow \sum w_j v_{ij}$$

When the items are then ranked according to this final score, items with high values for those criteria stated as being important by the user will appear towards the top of the ranked result set. Those items with lower values for the criteria stated as important will appear low in the ranked result set.

Technique for Order Preference by Similarity to Ideal Solution (TOPSIS)

The TOPSIS ranking method is based upon the concept that "*the chosen alternative should have the shortest distance from the ideal solution and the farthest from the negative-ideal solution.*" [20]

The two solutions are defined as follows:

- Ideal solution – a potential solution composed of all best attainable criterion values.
- Negative-ideal solution – a potential solution composed of all worst attainable criterion values.

By calculating the difference between each information item's quality score across all criteria, and the ideal and negative-ideal solutions, TOPSIS ranks each item according to how closely they match these two solutions: those closest to the ideal and furthest from the negative-ideal being ranked highly, and those furthest from the ideal and closest to the negative-ideal receiving a low ranking. The final order of the ranked set of data items will therefore be focused on those that best meet the stated needs of the user, based on their inputted weightings of importance for each criterion. The items ranked highly as a result of applying TOPSIS will be those which are the closest available match to the user's requirement and are therefore considered the 'best' items, and those ranked low will be those considered the 'worst' items against the user's requirement. The final set of ranked results are therefore focused according to the requirements of the user.

TOPSIS with Given Preferences (TOPSIS-GP)

The original version on TOPSIS does not take into consideration user preferences for ideal criteria values. We therefore created an updated version of this algorithm to incorporate user-specified ideal values for each quality criterion, rather than the highest (or lowest, for negative criteria) available values.

The current version of ISE only provides a facility for stating positive ideal values. Expanding ISE to allow the stating of negative ideal, or worst, values for each criterion would be a relatively small step. However, at present negative values are assumed based on the following:

If preferred ideal value for an IQ criterion is larger than the mean value, we assume the user prefers a high criterion value, so the negative-ideal is set to the lowest available value.

If preferred ideal value for an IQ criterion is lower than the mean value, we assume the user prefers a low criterion value, so the negative-ideal is set to the highest available value.

Using these assumptions the preference negative-ideal solution is created complementary to the preference positive ideal. Both are then used in the TOPSIS method instead of the standard ideal and negative-ideal, thus increasing search focus based on user-desired preferences.

Comparing Ranking Algorithms

Both the SAW and TOPSIS ranking algorithms were created to rank a set of data items into the best possible order, placing the 'best' items at the top of the ranking order and the 'worst' at the bottom. As can be seen however from the above discussions there are some major differences between the three algorithms. These differences are shown in Table 2.

Consideration	SAW	TOPSIS	TOPSIS-GP
Criterion preference weightings	✓	✓	✓
Positive and negative criteria[1]		✓	✓
Ideal values for criteria		✓	✓
User preference values for criteria			✓

Table 2 Comparing SAW, TOPSIS and TOPSIS-GP

[1] Criteria are considered positive if a high value is preferable over a low value (e.g. 'reliability'), and negative if a low value is preferable over a high value (e.g. 'price').

The comparison of quality criteria features taken into consideration by these algorithms shows that TOPSIS, and TOPSIS-GP, provide the opportunity for increased focusing of search results, to reflect the stated quality requirements of the user. By considering both positive and negative criteria, and ideal criterion values, TOPSIS ranks data according to how close each data item matches the user specified requirements across several criteria, based on criteria importance weightings. TOPSIS-GP increases ranking order focus by removing the assumption that the user is looking for the highest possible value for each criterion (or lowest for negative criteria), and ranks results according to both criteria importance weightings and how closely each criterion value is to the ideal value stated by the user.

Although these three algorithms are all available for use in ISE, TOPSIS-GP has been chosen as the primary algorithm due to its incorporation of user preferences. If no user preferences are stated then it defaults to the ideal values as used in the original TOPSIS algorithm.

EXPERIMENTS IN QUALITY FOCUSED INFORMATION SEARCHING

To evaluate the viability of using quality criteria to focus information searches a selection of simulations were conducted in the chosen subject domains. The purpose of these search simulations was to demonstrate the effects of changing quality preferences on the results obtained from a set of information searched, and to ascertain if any statistically significant differences occur in the ranking order of the results when quality preferences are changed.

Results of Simulations

The results presented in this section were produced using our experimental Information Search Environment (ISE). There simulations were conducted to discover whether the ranking order of search results can significantly change in the following situations:

- Changing the selection of quality criteria;
- Changing importance weightings for selected quality criteria;
- Changing preference values for selected quality criteria;
- Changing both importance weightings and preference values for selected quality criteria.

To ascertain the significance of the differences, if any, between the search results we used the Wilcoxon signed ranks test for statistical significance [10].

Below we discuss each of these four experiment types and conclude with a summary of our obtained results.

Simulation Set 1: Changing Selected Quality Criteria

The first set of simulations to be performed focused on the selection of individual quality criteria for each domain, and comparing the ranking order of the result set when the two searches were conducted. Two example results can be seen in Table 3. As the aim of simulation set 1 was not related to criteria importance weightings these were kept constant at 100%.

Domain	Quality Criteria	Weighting	Statement of Statistical Significance
University	Research Quality	100%	The differences noted in this simulation are very highly significant, to the 0.2% level.
	Facilities Spending	100%	
Cars	Drivability	100%	The differences noted in this simulation are very highly significant, to the 0.2% level.
	Safety	100%	

Table 3 Results of varying quality criteria selection

Simulations were also conducted into selecting distinct sets of quality criteria, as shown in Table 4.

Domain	Quality Criteria	Weighting	Statement of Statistical Significance
University	Dimension of Financial-Incoming	80%	The differences noted in this simulation are very highly significant, to the 0.2% level.
	Dimension of Veracity	80%	
Cars	Comfort, Drivability, Depreciation, Safety	100%	The differences noted in this simulation are very highly significant, to the 0.2% level.
	Reliability Running costs Value for money Window visibility	100%	

Table 4 Results of selecting distinct sets of quality criteria

Simulation Set 2: Changing Criteria Importance Weightings

The second set of simulations focused on the selection of two quality criteria for each domain, and comparing the results of using both criteria to focus a search as their importance weightings are changed. Two example results from this set are shown in Table 5.

Domain	Quality Criteria	Weighting	Statement of Statistical Significance
University	Teaching Quality Research Quality	10% 90%	The differences noted in this simulation are very highly significant, to the 0.2% level.
	Teaching Quality Research Quality	90% 10%	
Cars	Reliability Running costs	90% 10%	The differences noted in this simulation are very highly significant, to the 0.2% level.
	Reliability Running costs	10% 90%	

Table 5 Results of varying quality criteria importance values

Simulation Set 3: Changing Criteria Preference Values

The third set of simulations focused on the changing of quality criteria preferences vales, while keeping all other settings constant. Example results from this set are shown in Table 6.

Domain	Quality Criteria	Preference Value	Weighting	Statement of Statistical Significance
University	Research Quality	Max. possible value	100%	The differences noted in this simulation are very highly significant, to the 0.2% level.
	Research Quality	Min. possible value	100%	
Cars	Depreciation	Max. possible value	80%	The differences noted in this simulation are very highly significant, to the 0.2% level.
	Depreciation	Min. possible value	80%	

Table 6 Results of varying quality preference values, within identical criteria selection

Simulation Set 4: Changing Criteria Importance Weightings and Preference Values

The final set of simulations focused on changing both criteria importance weightings and preference values, while keeping the set of selected criteria constant. Example results from this set are shown in Table 7.

Domain	Quality Criteria	Preference Value	Weighting	Statement of Statistical Significance
University	Grad. Employment Dropout Rates	Max. possible value Min. possible value	100% 50%	The differences noted in this simulation are very highly significant, to the 0.2% level.
	Grad. Employment Dropout Rates	Min. possible value Max. possible value	50% 100%	
Cars	Drivability Reliability	Average value Max. possible value	50% 100%	The differences noted in this simulation are very highly significant, to the 0.2% level.
	Drivability Reliability	Max. possible value Average value	100% 50%	

Table 7 Results of varying importance weightings and preference values, with identical sets of quality criteria

Results of Simulations

The results presented above show that changing quality preferences can result in a highly statistically significant difference in the ranking order of the obtained results set.

However, we must also be aware that such noticeable differences are not always obtained. For example, if conducting a search based on a single quality criterion, changing the importance weighting of that criterion will not result in a different result set ordering, as the same ranking will be obtained if the criterion is rated at 100% or 50%. The quality scores for each item will change, but this change will be the same for all items, thus resulting in an identical ordering. Also, no change will be noted if several criteria are selected, but their importance weightings change in the same direction. For example, if while searching in the university domain the settings for the criteria *Research Quality* and *Teaching Quality* are changed from 100% and 80% to 50% and 40%, the ranking order will remain the same as importance values have not changed the relation between the selected criteria. Trivial changes in result ranking order are likely to be observed when making small changes to selected criteria preferences, such as when changing criteria preference values in minor steps.

Examples of situations in which little or no statistical significance is observed can be seen in Table 8.

Domain	Quality Criteria	Preference Value	Weighting	Statement of Statistical Significance
University	Teaching Quality	Max. possible value	100%	No statistical significance is seen in this example
	Gradate Employment	Max. possible value	100%	
Cars	Reliability	Max. possible value	100%	No statistical significance is seen in this example
	Performance Drivability	Max. possible value Max. possible value	90% 50%	
Cars	Drivability Performance	None stated	100% 80%	The differences noted in this simulation are significant, to the 0.1% level.
	Drivability Performance	None stated	80% 100%	

Table 8 Examples of no statistically significant difference in ranking order of results

The results presented in this paper therefore show that changes in the quality preferences of an individual consumer can result in a statistically significant difference in the ranking order of the result set, but that this is not necessarily always the case.

CONCLUSIONS

As stated in our research hypothesis, the aim of this project was to ascertain whether it was possible to produce a hierarchical generic model of quality that can be used by the information consumer to assist in information searching, by enabling them to state their quality preferences and using this information to focus information search results.

In previous papers we presented a user-oriented method for defining quality [7;8]. In this paper we have built on our previous work by showing how the information consumer can use quality criteria to focus information searches, within a number of experimental subject domains. By conducting a set of simulations, using quality criteria to focus information searches, we have shown that changing quality preferences can result in a statistically significant difference in the ranking order of the returned results.

Having shown that it is possible to develop a user-oriented model of quality, and that quality criteria can be used successfully to focus information search results, further work can now be done that builds on this project. Potential exists for research into a number of areas, including:

- Using quality preferences when searching for information in large or distributed environments;
- Filtering information based on quality preferences;
- User feedback, both implicit and explicit, to improve default quality settings;
- Automatic mapping of quality criteria to available data;
- Automatic relaxation of quality preferences.

The next phase of our work will therefore focus on investigating these potential areas, and others, to ascertain whether further development is needed, and in which direction to take this further work, while keeping in mind the importance of the information consumer on the quality lifecycle, as:

"*The customer or customers are the final arbiters of quality*" Redman (1996) [33].

ACKNOWLEDGEMENTS

We would like to thank EPSRC [16] for supporting M. Burgess with a studentship (award number 99302270), and BT [6] for providing supplementary funding. Thanks also go to Dr Wendy Ivins for providing encouragement and research-related advice.

BIBLIOGRAPHY

[1] Abate, M. L., Diegert, K. V., and Allen, H. W. "A Hierarchical Approach to Improving Data Quality," *Data Quality*, vol. 4, no. 1, 1998, pp.365-369.

[2] Aladwani, A. M. and Palvia, P. C. "Developing and Validating an Instrument for Measuring User-Perceived Web Quality," *Information and Management*, vol. 39, no. 6, 2002, pp.467-476.

[3] Barbacci, M., Klein, M. H., Longstaff, T. A., and Weinstoch, C. B. "Quality Attributes", CMU/SEI-95-TR-021 ESC-TR-95-021, Software Engineering Institute, Carnegie Mellon University, 1995.

[4] Boehm, B. W., Brown, J. R., and Lipow, M. "Quantitative Evaluation of Software Quality." *Proceedings of the 2nd International Conference on Software Engineering*, San Francisco, CA, USA, 1976, pp.592-605.

[5] Bovee, M., Srivastava, R. P., and Mak, B. "A Conceptual Framework and Belief-Function Approach to Assessing Overall Information Quality." *Proceedings of the 6th International Conference on Information Quality* (ICIQ-01), Cambridge, MA, USA, 2001, pp.311-324.

[6] BT. "*BT*" http://www.bt.com (Accessed on 1-5-2003).

[7] Burgess, M. S. E., Gray, W. A., and Fiddian, N. J. "Establishing a Taxonomy of Quality for Use in Information Filtering." *Proceedings of the 19th British National Conference on Databases* (BNCOD 19), Sheffield, UK, Springer, LNCS 2405, 2002, pp.103-113.

[8] Burgess, M. S. E., Gray, W. A., and Fiddian, N. J. "A Flexible Quality Framework for Use in Information Retrieval." *Proceedings of the 8th International Conference on Information Quality* (ICIQ-03), Cambridge, MA, USA, 2003, pp.297-313.

[9] Chen, Y., Zhu, Q., and Wang, N. "Query Processing with Quality Control in the World Wide Web," *World Wide Web*, vol. 1, no. 4, 1998, pp.241-255.

[10] Crichton, N. "Wilcoxon Signed Rank Test," *Journal of Clinical Nursing*, vol. 9, 2003, pp.584.

[11] Cykana, P., Paul, A., and Stern, M. "DOD Guidelines on Data Quality Management." *Proceedings of the Conference on Information Quality* (IQ-1996), Cambridge, MA, USA 1996, pp.154-171.

[12] Dedeke, A. "A Conceptual Framework for Developing Quality Measures for Information Systems." *Proceedings of the 2000 Conference on Information Quality* (IQ-2000), Cambridge, MA, USA, 2000, pp.126-128.

[13] Dromey, R. G. "A model for software product quality," *IEEE Transactions on Software Engineering*, vol. 21, no. 2, 1995, pp.146-162.

[14] Dromey, R. G. "Concerning the Chimera," *IEEE Software*, vol. 13, no. 1, 1996, pp.33-43.

[15] Eppler, M. J. "A Generic Framework for Information Quality in Knowledge-Intensive Processes." *Proceedings of the 6th International Conference on Information Quality* (IQ 2001), Cambridge, MA, USA, 2001, pp.329-346.

[16] EPSRC. "*EPSRC*" http://www.epsrc.ac.uk (Accessed on 1-5-2003).

[17] Euro NCAP. "*European New Car Assessment Programme*" http://www.euroncap.com (Accessed on 1-8-2003).

[18] Gardyn, E. "A Data Quality Handbook for a Data Warehouse." *Proceedings of the 1997 Conference on Information Quality* (IQ-1997), Cambridge, MA, USA, 1997, pp.267-290.

[19] Guardian, The. "*The Guardian Home Page*" http://www.guardian.co.uk (Accessed on 4-6-2003).

[20] Hwang, C. L. and Yoon, K. *Multiple Attribute Decision Making: Methods and Applications*, Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, New York, 1981.

[21] Hyatt, L. E. and Rosenberg, L. H. "A Software Quality Model for Identifying Project Risks and Assessing Software Quality." *Proceedings of the 8th Annual Software Technology Conference*, Utah, USA, 1996.

[22] ISO/IEC 9126-1:2001 "Information Technology - Software Product Quality", 2001.

[23] Liu, K., Zhou, S., and Yang, H. "Quality Metrics of Object Oriented Design for Software Development and Re-Development." *Proceedings of the 1st Asian Pacific Conference on Quality Software* , Hong Kong, China, 2000, pp.127-135.

[24] Long, J. A. and Seko, C. E. "A New Method for Database Data Quality Evaluation at the Canadian Institute for Health Information (CIHI)." *Proceedings of the 7th International Conference on Information Quality* (ICIQ-02), Cambridge, MA, USA, 2002, pp.238-250.

[25] Matsumura, A. and Shouraboura, N. "Competing with Quality Information." *Proceedings of the 1996 Conference on Information Quality* (IQ-1996), Cambridge, MA, USA, 1996, pp.72-86.

[26] McCall, J. A., Richards, P. K., and Walters, G. F. *Factors in Software Quality, Volumes I-III,* National Technical Information Service, 1977.

[27] Miller, H. "The Multiple Dimensions of Information Quality," *Information Systems Management*, vol. 13, no. 2, 1996, pp.79-82.

[28] Naumann, F. *Quality-Driven Query Answering for Integrated Information Systems,* Springer-Verlag, 2002.

[29] Naumann, F. and Rolker, C. "Assessment Methods for Information Quality Criteria." *Proceedings of the 2000 Conference on Information Quality* (IQ-2000), Cambridge, MA, USA, 2000, pp.148-162.

[30] Olsina, L., Lafuente, G., and Rossi, G. "Specifying Quality Characteristics and Attributes for Web Sites," in Murugesan, S. and Deshpande, Y. (eds.) *Web Engineering 2000,* Springer-Verlag, Berlin, Germany, 2001, pp.266-278.

[31] Ortega, M., Pérez, M. A., and Rojas, T. "A Systemic Quality Model for Evaluating Software Products." *Proceedings of the 8th World Multi-Conference on Systemics, Cybernetics and Informatics* (SCI 2002/ISAS 2002), 2002.

[32] Parkers. "*Parkers Web Site*" http://www.parkers.co.uk/ (Accessed on 15-8-2003).

[33] Redman, T. *Data Quality for the Information Age,* Artech House, MA, USA, 1996.

[34] Royce, W. "Pragmatic Quality Metrics for Evolutionary Software Development Models." *Proceedings of the Conference on TRI-ADA '90*, Baltimore, Maryland, USA, ACM Press, 1990, pp.551-565.

[35] Rubey, R. J. and Hartwick, R. D. "Quantitative Measurement of Program Quality." *Proceedings of the ACM National Conference* 1968, pp.671-677.

[36] Smart, K. L. "Assessing Quality Documents," *ACM Journal of Computer Documentation*, vol. 26, no. 3, 2002, pp.130-140.

[37] Times, The. "*The Times Online Newspaper*" http://www.timesonline.co.uk/ (Accessed on 4-6-2003).

[38] Wang, R. Y. and Strong, D. M. "Beyond Accuracy: What Data Quality Means to Data Consumers," *Journal of Management Information Systems*, vol. 12, no. 4, 1996, pp.5-34.

[39] What Car. "*What Car Home Page* " http://www.whatcar.com/ (Accessed on 15-8-2003).

[40] Which? Online. "*Which? Online Home Page*" http://www.which.net/ (Accessed on 4-6-2003).

[41] Zhu, X. and Gauch, S. "Incorporating Quality Metrics in Centralized/Distributed Information Retrieval on the World Wide Web." *Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (SIGIR-00), Athens, Greece, 2000, pp.288-295.