International Journal of Computational Intelligence Systems, Vol.3, No. 4 (October, 2010), 420-437

A Review on Information Accessing Systems Based on Fuzzy Linguistic Modelling

E. Herrera-Viedma, A.G. López-Herrera

Department of Computer Science and Artificial Intelligence,
CITIC-UGR (Research Center on Information and Communications Technology),
University of Granada, E-18071-Granada, Spain
E-mails: {viedma, lopez-herrera}@decsai.ugr.es

Received: 13-02_2010
Accepted: 25-05-2010

Abstract

This paper presents a survey of some fuzzy linguistic information access systems. The review shows information retrieval systems, filtering systems, recommender systems, and web quality evaluation tools, which are based on tools of fuzzy linguistic modelling. The fuzzy linguistic modelling allows us to represent and manage the subjectivity, vagueness and imprecision that is intrinsic and characteristic of the processes of information searching, and, in such a way, the developed systems allow users the access to quality information in a flexible and user-adapted way.

Keywords: Fuzzy linguistic modelling, information access systems, information retrieval systems, recomender systems, quality information.

1. Introduction

With the growth of Internet, many on-line information systems are coming up, and the amount of information available makes necessary the development and use of effective Information Access Systems (IASs) that allow user easy and flexible access to quality and relevant information.

These IASs must be user-adapted by two reasons:

- On the one hand, information seeking, retrieval and filtering are inherent human abilities which are not necessarily rationally guided, which can be based on explicit or tacit assumptions and that they do not need precise and complete measurements about the set of feasible information items (usually documents); and
- On the other hand, the concepts of "relevance" and "quality" are two key concepts in IASs, however, only the user can determine the real relevance or actual quality of an information item, i.e., the usefulness, pertinence, appropriateness, or utility of that information item with respect to his/her information needs and/or preferences (usually expressed as a user query or user profile).

Soft Computing (SC) [1] constitutes a synergy of methodologies (including fuzzy set theory, neural networks, probabilistic reasoning, rought-sets, evolutionary computing and some approaches of machine learning) which are useful for solving problems requiring some form of "intelligence". The advantage of using SC is its tolerance to imprecision, uncertainty, partial truth, and approximation. Due to these properties, techniques based on SC are very suitable for modelling the activities related to the information access problem (for more information see

[2,3,4,5,6,7,8]).

The fuzzy linguistic modelling (FLM) is a SC tool very useful to represent qualitative information in the problems. FLM is based on the concept of "linguistic variable" [9,10,11] which was introduced by the Prof. Zadeh to model those variables whose values are words or sentences (not numbers) in a natural or artificial language. The main purpose of using linguistic values instead of numbers is that linguistic characterizations are, in general, less specific than numerical ones, but much closer to the way that humans express and use their knowledge. For example, if we say the building is tall, this sentence is less specific than the building measures 300 m. In that case, tall can be seen as a linguistic value of the variable height which is less precise and informative than the numerical value 300. Despite its less informative nature, the value tall allows humans to naturally express and deal with information that may be uncertain or incomplete (the speaker may not know the exact building height)[12]. As this kind of situations where information is not precise is very common in real life, linguistic variables can be a powerful tool to model human knowledge (see [13,14,15,16,17,18,19,20]). Some of the most important FLM approaches to model linguistic information are:

- *Classical FLM* [21,22,23]*:* This approach makes use of membership functions to model and combine linguistic assessments.
- *Type-2 FLM* [24,25,26,27]*:* This approach makes use of type-2 fuzzy sets to model and combine linguistic assessments.
- *Symbolic FLM* [28,29,30]*:* This approach uses ordered and symmetrically distributed linguistic term sets with odd cardinality and then aggregation is made directly acting over the label indexes.
- *2-tuple FLM* [31]*:* This is a symbolic approach that improves the previous one by representing the linguistic information by means of a pair of values called *linguistic 2-tuple* (s, α), where s is a linguistic term and α is a numeric value representing a *Symbolic Translation*.
- *Muli-granular FLM* [32,33,34,35]*:* This approach assumes that in many problems is necessary to use different linguistic term sets with different semantics and cardinalities to model the linguistic assessments.
- *Unbalanced FLM* [36,37,38]*:* This approach assumes that in real situations we need to work with non-symmetrical linguistic term sets.

This paper presents an overview of IASs based on FLM focusing on three technologies related to the information access process:

1. Information retrieval systems [39,40,41],
2. Filtering and recommender systems [42,43,44], and
3. Web quality evaluation tools [45,46,47].

We analyze their performance and how such FLM based technologies allow users the access to quality and relevant information in a best flexible and user-adapted way.

To do this, the paper is organized as follows: Section 2 briefly introduces the technologies related to IASs that we revise in this paper. Section 3 analyzes the performance of such FLM based technologies. Finally, some conclusions are drawn in Section 4.

2. On Technologies Related to Information Access Systems

This section introduces the main technologies related to the IASs that have shown to be a useful application field for the FLM: information retrieval systems, filtering and recommender systems, and models for web quality evaluation.

2.1. Information retrieval systems

Information retrieval (IR) may be defined as the problem of the selection of documentary information from storage in response to search questions provided by a user [39,48]. Information Retrieval Systems (IRSs) deal with documentary bases containing textual, pictorial or vocal information and process user queries trying to allow the user to access to relevant information in an appropriate time interval. Both documents and user queries must be formally represented in a consistent way, so that IRSs can satisfactorily develop the retrieval activity. Nowadays, the development of the WWW has increased the interest on the study of IRSs.

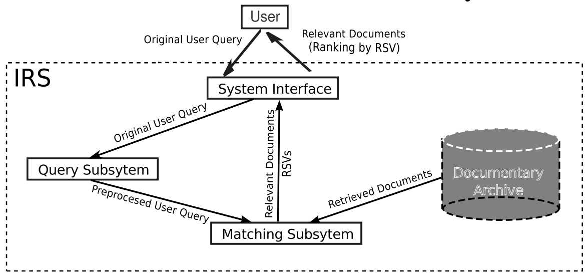

Fig. 1. Components of an IRS

An IRS is basically constituted by three main components (see Figure 1):

1. *A documentary archive* which stores the documents and the representation of their information contents (index terms). It is built using tools for extracting index terms and for representing the documents.

2. *A query subsystem* which allows users to formulate their information needs (queries) by means of a formal query language.

3. *A query evaluation component* which evaluates the documents for a user query obtaining. So, it presents an inference procedure that establishes a relationship between the user request and the documents stored in the documentary archive in order to determine the relevance of each document to the user query.

Most of the existing IRSs are based on the Boolean retrieval model [49]. Usually, in the documentary archive the documents are represented as sets of index terms, the query component represents the user queries as Boolean combinations of index terms, and the evaluation component uses a total matching mechanism between documents and queries as an inference procedure. These IRSs present many limitations [40], mainly the lack of flexibility and precision for representing document contents, for describing user queries and for characterizing the relevance of the documents retrieved for a given user query. These drawbacks may be overcome by incorporating weights in the three levels of information representation of an IRS:

1. *Document representation level.* By computing weights of index terms, the system specifies to what extent a document matches the concept expressed by the index terms.

2. *Query representation level.* By attaching weights in a query, a user can provide a more precise description of his or her information needs or desired documents.

3. *Evaluation representation level.* By assigning weights to characterize the relationships between user queries and document representations the evaluation subsystem provides a means, called retrieval status value (RSV) of a document, in order to discriminate the documents retrieved by relevance judgments.

Fuzzy Set Theory [50] has been used in order to achieve a mathematical formalization of the use of weights for handling uncertain information in all information representation levels of an IRS [51,52,53,54,55,56]. Particularly, we should point out that we can find in the literature some fuzzy IRSs enriched with weighted query languages [57,58,59,60] that increase the expressiveness of the traditional Boolean query languages [49], allow users represent better in the queries their concept of relevance, and improve the effectiveness of IRSs. Furthermore, they provide different semantics associated with the weights of the queries [58]:

1. *Importance semantics*, considering the weights as measures of the importance of a specific element in representing the query.

2. *Threshold semantics*, considering the weights as a threshold to aid in matching a specific document to the query.

3. *Perfection semantics*, considering the weights as a description of an ideal or perfect document.

These fuzzy IRSs use predominantly numeric weights (values in [0,1]) to weigh the user queries. This limits the user expressiveness, and therefore, they should be able to take into account the possibility for using qualitative values typical of human communication. To do so, the FLM has been applied satisfactorily by allowing to define new linguistic weighted query languages. In the Section 3.1 a survey of IRSs based on FLM is presented.

2.2. Filtering or recommender systems

Information gathering in Internet is a complex activity. Find the appropriate information, required for the users, on the Web is not a simple task. To improve the information access on the Web the users need tools to filter the great amount of information available to assist users in the information gathering process to access to quality information in a user-adapted way. Filtering systems or recommender systems (RSs) offer tools for discriminating between relevant and irrelevant information by providing personalized assistance for continuous information accesses, filtering the information and delivering it to people who need it [43].

RSs to filter information can be characterized by the following aspects [42,43,61]:

- They are applicable for unstructured or semi-structured data (e.g. Web documents or e-mail messages).
- Users have long time information needs that are described by means of user profiles.
- They handle large amounts of data.
- They deal primarily with textual data.
- Their objective is to remove irrelevant data from incoming streams of data items.

We can find some of the above features in IRSs, but a RS differs from traditional IRS in that the users have long information needs that are described by means of user profiles, rather than ad-hoc needs that are expressed as queries posed to some IRS. Traditionally, an IRS develops storage, indexing and retrieval technology for textual documents. A user describes his information need in the form of a query to the IRS and the system attempts to find items that match the query within a document store. The information needs are usually very dynamic and temporary, i.e., a user uses a query describing an immediate need. Furthermore, IRSs tend to maintain a relatively static store of information. Unlike IR systems, RSs generally operate on continuous information streams, and always maintain a profile of the user interests needs throughout many uses of the system. As a result, Rss tend to filter information based on more long-term interests.

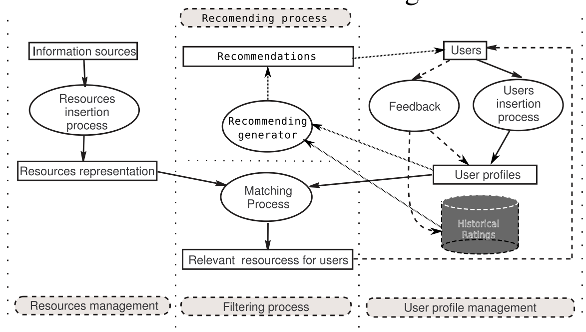

Fig. 2. Components of a RS

The classical key components of a RS to filter information are [42,43,61] (see Figure 2):

1. *The user profiles* to represent long time user information needs.
2. *The representation of information items or documents.*
3. *The method of generating recommendations.*
4. *The set of historic ratings* provided by the users that received recommendations.

The construction of accurate profiles is a key task and the system's success will depend on a large extent on the ability of the learned profiles to represent the user's preferences [62] or user information. In order to discriminate between relevant and irrelevant information for a user, we must have some information about this user, i.e. we must know the user preferences. Information about user preferences can be obtained in two different ways [42,62], *implicit* and *explicit mode*, although these ways do not be mutually exclusive. The implicit approach is implemented by inference from some kind of observation.

The observation is applied to user behavior or to detecting a user's environment (such as bookmarks or visited URL). The user preferences are updated by detecting changes while observing the user. On the other hand, the explicit approach, interacts with the users by acquiring feedback on information that is filtered, that is, the user expresses some specifications of what they desire. Two desired properties that any user profiling should support are the following:

1. User profiles should be adaptable or dynamic since user's interests are changing continuously and rapidly over time. This implies the need to include a learning module in the RS to adapt the user profile according to feedback from user reaction to information provided by the system.

2. The generation and update of user profiles should be carried out with a minimal explicit involvement of the users, i.e. by minimizing the degree of the user intervention to reduce user effort and facilitate the system-user interaction.

On the other hand, depending on the method of generating recommendations, traditionally RSs have been grouped into two main categories [42,43]:

- *Content-based RSs* which filter and recommend the information to an active user by matching terms used in the representation of documents and the ratings that the active user has previously given to them ignoring data from other users. In a content-based RS, the recommendation is done by the system itself, i.e., the function of the system is to provide recommendations for its individual users, as, for example, the order of the documents in a given collection. In this sense, a recommender system is almost like an IRS [63]. These RSs tend to fail when little is known about active user's preferences.
- *Collaborative RSs* which use explicit or implicit preferences from many users to filter and recommend documents to a given user, ignoring the representation of documents. Collaborative systems locate peer users with user profiles and/or rating history similar to the current user and they generate recommendations using this neighborhood. On the other hand, in a collaborative recommender system, the recommending is done by the users of system, i.e., the function of the system is to synthesize multiple users recommendations of documents in the form of a single ranking for the individual user [63]. These RSs tend to fail when he/she has uncommon interests.

All of these techniques have benefits and disadvantages. However, we can use a hybrid approach to smooth out the disadvantages of each one of them and to exploit their benefits. We should point out that the generation process of recommendations involves two steps [64]:

1. *Computation of similarity degrees*. In the case of content-based RSs, the similarity degrees are calculated between a new unexperienced document and other documents that user has experienced and rated previously. In the case of collaborative RSs, the similarity degrees are calculated between our user profile and other user profiles, without considering the document representations.

2. *Aggregation of ratings*. In the case of content-based RSs, the ratings are supplied by the user who receives the recommendation. In the case of collaborative RSs, the ratings are supplied by other users.

The recommendation activity is followed by a relevance feedback phase. *Relevance feedback* is a cyclic process whereby the user feeds back into the system ratings on the relevance of retrieved documents and the system then uses these evaluations to automatically update the user profile [42,43,65].

Fuzzy Set Theory [50] has been applied satisfactorily in RSs to manage the uncertainty in the representation of user profiles and in the generation process of recommendations [64,66]. In Section 3.2 some RSs based on FLM are revised .

2.3. Models for web quality evaluation

Nowadays, everybody knows that the Internet is the largest available repository of data with the largest number of visitors searching for information. However, its growth fast, disorganized and uncontrolled, its heterogeneity, lack of publishing control, have contributed to that bad information thrives on the World Wide Web [67]. As a consequence Internet users have access to bad or poor-quality information and this problem would be solved by using mechanisms for filtering low-quality information on the Internet [68].

There exists much debate on the quality of the information available on the Web, and how to recognize useful and quality information in an unregulated market place such as the Internet [67]. Unfortunately, in the literature, one can probably find as many definitions for information quality on the Web as there are papers on information quality. Due to the quality evaluation on the World Wide Web is neither simple nor straight forward, and as consequence, there is not a general theoretical foundation or framework in Web quality evaluation [69].

So, many researchers have tried to use other well-founded quality assessment frameworks defined for other fields. One of the more often used is the information quality framework defined in the context of management information systems [70,71,72,73]. This quality framework establishes that the different dimensions (e.g., accuracy, accessibility, relevance) employed to evaluate the information quality of a system can be grouped into four major information quality categories: (1) intrinsic information quality, (2) contextual information quality, (3) representational information quality, and (4) accessibility information quality. The two first information quality categories mainly deal with the "content" aspects of information systems, the others, with some technical design aspects.

The evaluation of Web sites focusing on the quality of the information that it provides is a difficult task that has rarely been studied [47]. However, a robust and flexible Web quality evaluation methodology should properly combine both kinds of requirements, content and technical ones. Some authors [74,75,76] have proposed Web quality evaluation methodologies that combine both technical and content aspects, but the harsh reality is that the majority of suggested Web evaluation methodologies tend to be more objective than subjective, more quantitative than qualitative, and do not take into account the user perception [46,77,78]. However, from the information consumer's perspective the quality of a Web document/site may not be assessed independently of the quality of the information contents that it provides [79]. An additional drawback of many Web evaluation methodologies is that their evaluation indicators are relevant to Web providers and designers rather than to Web users [45]. A global Web quality evaluation methodology cannot entirely avoid users' participation in the evaluation strategy.

Usually, a global Web quality evaluation methodology presents the following general components [75,79,80]:

- *An evaluation scheme*: It establishes the different evaluation criteria or indicators to be considered in the evaluation of Web resources and their importance degrees. Usually, it is appropriate to take into account both subjective and objective criteria and the users participation.
- *A measurement method*: It establishes how to obtain the ratings associated with each evaluation criteria (e.g., we would have to define a questionnaire to gather users' perceptions) and an aggregation or synthesis mechanism to obtain the quality rating associated with the particular Web resource (site or document).

In many commercial information systems (specially in those that manage some kind of recommendations) is usual users provide evaluation judgements or annotations about products as inputs, which the system then aggregates obtaining recommendations that are stored (see for example the recommendations polities used by Amazon[*] TripAdvisor[†]or Booking[‡]). Then, these recommendations can

[*] http://www.amazon.com/
[†] http://www.tripadvisor.es/
[‡] http://www.booking.com/

be used to assist other users in their search process for localizing similar products. In this sense, recommendations are a kind of plausible measure of the quality of those products. From the point of view of an user demanding quality information, judgments can help to evaluate the information quality of accessed Web documents/sites because the concept of information quality is typically consumer dependent, and the consumer must be the ultimate judge of the Web site's/document's information quality.

The problem here is that on the one side, the users do not frequently make the effort to give explicit feedback, on the other side, systems tend to force users to give theirs evaluations in a very strict way, usually in form of numerical values [43], but sometimes, a person cannot express his/her judgements with an exact numerical value. A possible way to facilitate the user participation is to embed in the Web quality evaluation methodology those tools of Artificial Intelligence that allow a better representation of subjective and qualitative user judgments, for example, the FLM [9,10,11]. The use of FLM could increase user participation in the evaluation of the quality of Web documents/sites, because it is a user-friendly tool that helps users to express their judgments in a more natural way using words rather numerical values [79]. Several Web quality evaluation methodologies using FLM have been proposed, some of them for evaluating the quality of Web documents (HTML, XML, SGML, RSS, etc.), and others focused on evaluating the quality of entire Web sites. Section 3.3 reviews some papers approaching the quality evaluation of Web resources (documents or sites) by means of FLM approaches.

3. Using Fuzzy Linguistic Modelling in Information Access Systems

This section reviews IASs which were designed based on FLM. The review is mainly focused on information retrieval systems, filtering and recommender systems and models of web quality evaluation.

3.1. Information retrieval systems based on fuzzy linguistic modelling

As aforementioned, the fuzzy IRSs that use weighted query languages based on numeric weights force the user to quantify qualitative concepts (such as "importance"), ignoring the fact that many users are not able to provide their information needs precisely in a quantitative form but in a qualitative one. In fact, it seems more natural to characterize the contents of the desired documents by explicitly associating a linguistic descriptor to a term in a query, such as "important" or "very important", instead of a numerical value. Similarly, IRSs are more user-friendly if the estimated relevance levels of the documents are supplied in a linguistic form (e.g., linguistic terms such as "relevant", "very relevant" may be used) rather than with scores. Following these ideas, several fuzzy linguistic IRSs have been proposed using a FLM to model the weighted user queries and the system output.

We identify the following fuzzy linguistic IRSs approaches:

1. Fuzzy linguistic IRSs based on classical FLM.
2. Fuzzy linguistic IRSs based on symbolic FLM.
3. Fuzzy linguistic IRSs based on 2-tuple FLM.
4. Fuzzy linguistic IRSs based on multi-granular FLM.
5. Fuzzy linguistic IRSs based on unbalanced FLM.

3.1.1. Fuzzy linguistic IRSs based on classical FLM

In the literature we can find three fuzzy linguistic IRSs based on a classical FLM: [81], [82] and [83].

In [81] Gloria Bordogna and Gabriella Pasi present the first fuzzy linguistic IRS which is based on a classical FLM. They define an extended Boolean IRS that supports weighted user queries and retrieves documents assessed by means of linguistic

RSVs. They define the linguistic variables ”Importance” and ”Relevance” by means of a context-free grammar to assess the query weights and the RSVs, respectively. The weights could be assigned on the queries terms of the queries and their interpretation was done by means of a perfection semantics. Therefore, in this system the linguistic weights are introduced by users to specify his/her concept of “ideal” documents [58].

In [82] Donald H. Kraft, Gloria Bordogna and Gabriella Pasi redefine the previous fuzzy linguistic IRS [81] by introducing a new threshold semantics in a linguistic context which was obtained by combination of both, the ideal semantics [58] and threshold semantics [60].

The last approach of fuzzy linguistic IRS based on a classical FLM was defined in [83]. Gloria Bordogna and Gabriella Pasi present a new fuzzy linguistic IRS assuming linguistic weighted queries with a perfection semantics. The main novelty of this system is that a new fuzzy representation of documents is defined in which different degrees of significance are computed for a given term, one for each document subpart, based on the subparts' semantics. Using this new document representation, which is dynamically interpretable by the user, a same weighted query can select documents in different relevance orders, depending on both the subparts' preferences and the aggregation criterion specified by the user by means of linguistic quantifiers [84].

3.1.2. *Fuzzy linguistic IRSs based on symbolic FLM.*

The fuzzy linguistic IRSs based on a symbolic FLM that we can find in the literature are the following: [17], [85] and [86].

Enrique Herrera-Viedma presents in [85] a fuzzy linguistic IRSs that defines the linguistic variables ”Importance” and ”Relevance” by means of a symbolic FLM. This IRS allows users to use multi-weighted queries to represent his information needs. The terms of a Boolean query can be weighted by linguistic weights according to three different semantics, threshold semantics, importance semantics, and a quantitative semantics (this semantics provides restrictions on the number of documents to be retrieved for each term), which could be used simultaneously o not according to the user preferences. The system includes a new definition of the threshold semantics, called symmetrical threshold semantics, which deals with an inverse interpretation to the linguistic weights to the left of the middle linguistic term and the weights to the right. It also introduces a new mechanism to evaluate the fuzzy linguistic multi-weighted Boolean queries that supports the consistency among the different semantics of the weights and uses the linguistic version of t-norm MIN and t-conorm MAX to model the connectives AND and OR in the evaluation of queries, respectively.

In [17] Enrique Herrera-Viedma proposes an extension of the previous symbolic linguistic IRS [85] that uses also multi-weighted queries based on two weighting elements: the query terms and the query sub-expressions. In such a way, users may easily express simultaneously several semantic restrictions in a query as in [17]. A symmetrical threshold semantics is associated to the weights of the query terms and an importance semantics is associated to the weights of the query sub-expressions. The first one is modelled by a linguistic matching function that is easier than that proposed in [17]. The latter is modelled by means of two aggregation operators of weighted linguistic information [28], the Linguistic Weighted Disjunction operator and the Linguistic Weighted Conjunction operator, used to model the connectives AND and OR in the subexpressions, respectively.

Both fuzzy linguistic IRSs, [17] and [85], are the basis of two important proposals that we have implemented:

1. All the previous FLM based IRSs are very powerful, allowing users to express theirs information needs in a very flexible and user-adapted way. However, there are situations (specially with non-expert users) in which users do not know or can not express their information needs directly by means of a simple or weighted Boolean query or weighted. Then, to overcome this problem we could apply automatic aid tools, as the *Multi-Objective Evolutionary Algorithms* (MOEAs) [87], to help

users to build those queries that better could represent their information needs. In [88] we present a MOEAs based automatic aid tool to help users in the building of fuzzy linguistic Boolean queries based on a symbolic FLM.

2. An educational software tool to teach fuzzy linguistic IRSs was implemented in [89]. The main purpose of this tool is to assist students and non-expert users in the complex process of learning the performance of the fuzzy weighted IRSs based on FLM. This tool allows students to compare the performance of different fuzzy IRSs proposed by other authors, including those based on numerical weights and those based on FLM. With such tool, student and non-expert users can use different weighted semantics (classical threshold or symmetrical threshold or relative importance or perfection or quantitative) on the queries and see how they are evaluated in the IRS. Two different FLM approaches can be used: symbolic and 2-tuple one.

On the other hand, Gloria Bordogna and Gabriella Pasi also propose a symbolic linguistic IRS in [86]. Linguistic expressions were defined to represent and manage the importance of both the index terms as descriptors of the information items and the query terms as descriptors of user information needs. In this IR model three weighting semantics (relative importance, threshold and ideal significance) can be used in user's queries, and quantifier based OWA operators are used in the evaluation process of queries to model the logical connectives of the queries.

3.1.3. Fuzzy linguistic IRSs based on 2-tuple FLM

The fuzzy linguistic IRSs based on a 2-tuple FLM that have been defined are the following: [90], [91], and [92].

A multi-agent system for IR purposes in the Internet was proposed in [90] by Miguel Delgado et al. They present a distributed intelligent agent model where the communication of the evaluation of the retrieved information among the agents is carried out by using linguistic information assessed on a symbolic linguistic approach, but the representation of linguistic information inside system is based on the 2-tuple fuzzy linguistic representation model and the computation of the RSV of the documents on the 2-tuple computational model.

In [91] Enrique Herrera-Viedma and Antonio G. López-Herrera and Carlos Porcel propose a fuzzy linguistic IRS based on the 2-tuple FLM that supports weighted queries based on a new interpretation of the symmetrical threshold semantics defined in [85]. The use of the 2-tuple FLM allows defining a new matching functions that improves the interpretation of the symmetrical threshold semantics proposed in [85].

Enrique Herrera-Viedma et al. present in [92] a new fuzzy linguistic IRSs based in a 2-tuple FLM that extended that IRS defined in [85] and solved the detected problems: i) loss of information, ii) loss of precision, and iii) rigid interpretation of the Boolean connectives optimistic evaluation of the satisfaction of the threshold values. The application of the 2-tuple FLM solves the two former problems, and the latter one was solved by introducing a new soft computing operator to model the Boolean connectives in a more flexible way, the 2-tuple linguistic LOWA (Linguistic Ordered Weighted Averaging) operator.

3.1.4. Fuzzy linguistic IRSs based on multi-granular FLM

In the contribution [93] Enrique Herrera-Viedma et al. propose an IRS based on a multi-granular FLM [32]. In this IRS we assume that in the activity of an IRS, there are aspects of different nature to be assessed, e.g., the relevance of documents, the importance of query terms, etc. Therefore, these aspects should be assessed with different uncertainty degrees, i.e., using several label sets with different granularity of uncertainty. Therefore, this new system accepts multi-weighted Boolean queries whose terms can be simultaneously weighted by means of ordinal linguistic values according to three semantics as in [85]: a symmetrical threshold semantics, a relative importance semantics and a quantitative semantics. But, in this case, each semantics is associated with a different label set S^1, S^2 and S^3, respectively, which could have a different granularity. Furthermore, the IRS evaluates multi-weighted queries and obtains the linguistic RSVs of documents represented by the linguistic variable Relevance which is also expressed on a different label set S'.

3.1.5. Fuzzy linguistic IRSs based on unbalanced FLM

Many fuzzy linguistic IRSs [17,85,86,91,92] uses symmetrically and uniformly distributed linguistic term sets in their retrieval activity (see Figure 3). In these cases, the same discrimination levels on both sides of the middle linguistic term are established. However, usually users look for documents with positive criteria, that is, they formulate their weighted queries using linguistic assessments on the right of the middle label much more than on the left. Similarly, usually users are interested in the relevant documents much more than in the non-relevant documents, and then, a best tuning of the output of the IRS can be achieved if a higher number of discrimination levels on the right of the middle linguistic term is assumed (see Figure 4). So, Enrique Herrera-Viedma and Antonio G. López-Herrera propose in [38] the first model of IRS based on an unbalanced FLM. This new unbalanced linguistic IRS accepts multi-weighted queries whose weights are expressed using unbalanced linguistic term sets and interpreted according to an importance semantics and threshold semantics. Then, the system provides the retrieved documents classified in linguistic relevance classes assessed on unbalanced linguistic term sets. To do so, we defined a first approach of a methodology to manage unbalanced linguistic information.

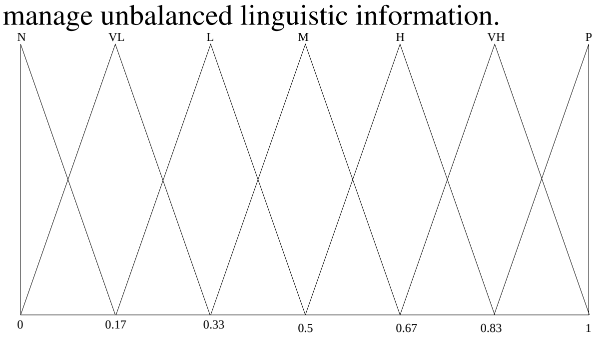

Fig. 3. Symmetrically Distributed Linguistic Term Set

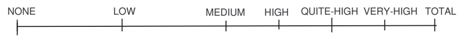

Fig. 4. Unbalanced Linguistic Term Set

3.2. Recommender systems based on fuzzy linguistic modelling

As aforementioned, the goal of RSs related to the information access processes is to evaluate and filter the great amount of information available on the Web or in particular contexts (e.g. academic libraries) to assist users in the continuous information gathering process to access to quality and relevant information. While an IRS responds to current user information needs, a RS tries to predict future user information needs. The success of the RS activity depends very much on the tools used to characterize and update the user profiles and generate the recommendations to be sent to the users. Fuzzy Set Theory [50] has been applied satisfactorily in RSs to manage the uncertainty in the representation of user profiles and in the generation process of recommendations [66,64]. In this section, we analyze different RSs approaches designed using FLM:

- Fuzzy linguistic RSs based on Symbolic FLM.
- Fuzzy linguistic RSs based on 2-Tuple FLM.
- Fuzzy linguistic RSs based on Multi-granular FLM.

3.2.1. Fuzzy linguistic RSs based on symbolic FLM

Three main fuzzy linguistic RSs based on a Symbolic FLM have been defined, [94], [95] and [96].

Zheng Pei et al. [94] extend the multi-agent IRS proposed in [90] but using a Symbolic FLM. As main novelties they present new extensions of linguistic OWA operators defined in [28,29], which allow for gathering and filtering Web documents from the linguistic weighted Boolean queries provided by the users.

In the paper by José M. Morales-del-Castillo et al. [95] a particular fuzzy linguistic hybrid RS based on a Symbolic FLM for digital libraries [97] specialized in Library and Information Sciences, called multi-agent system of Selective Dissemination of Information (SDI), is presented. It combines multi-agent technologies as those presented in [90], semantic Web technologies [98] (used to define rich descriptions of resources and a conceptual scheme that helps in indexing and retrieving tasks) and FLM tools [28,29] in order to generate and disseminate personalized bibliographic alerts for the users in a digital library.

By continuing with their research work, José M. Morales del Castillo et al. define in [96] an extension of fuzzy linguistic RS [99]. In this case, this new fuzzy linguistic RS is specialized in Biomedical Sciences. The main aim is to help biomedical workers to be updated in the most relevant biomedical publications retrieved from large biomedical repositories according to their user profiles that are generated automatically from a partial expression of their information needs. There are special situations, as for example, when the information related to the users and items is scarce and insufficient, where classical RSs (collaborative and content-based ones) have many problems to make good recommendations. Therefore, to increase the knowledge on the users the main novelty of this hybrid RS is the definition of a procedure to build automatically the user profiles from the information needs expressed by the users by means of incomplete fuzzy linguistic preference relations [100,101].

3.2.2. *Fuzzy linguistic RSs based on 2-tuple FLM*

We can find in the literature two fuzzy linguistic RSs based on a 2-tuple FLM: [99] and [102].

In [102] E. Herrera-Viedma et al. present a fuzzy linguistic RS model which incorporates information filtering possibilities in the multi-agent IRS defined in [90], but now exploiting the advantages of a 2-tuple FLM. This RS is based on a fuzzy linguistic multi-agent model for information gathering on the Web that implements content-based and collaborative information filtering techniques to improve the retrieval process. Users specify their information needs by means of both a linguistic multi-weighted query and an information need categories. In the multi-weighted queries two query weights can be simultaneously used. On the one hand, threshold weights are used by the content-based filtering agents to carry out a first filtering of documents to retrieve, and on the other hand, relative importance weights are used by the task agent to determinate the number of documents to be retrieved from each content-based filtering agent. The multi-agent model incorporates in its architecture a collaborative filtering agent that filters and recommends documents related to information need category according to the linguistic evaluation judgements previously expressed by other users.

José M. Morales del Castillo et al. present in [99] a new fuzzy linguistic RS based on a 2-tuple FLM, called D-Fussion, that improved their previous fuzzy linguistic SDI service model [95]. This new multi-agent system incorporates a new recommendation approach to satisfy researchers' specific information requirements that generates two kinds of linguistic recommendations, mono-disciplinary or specialization bibliographic recommendations (which are oriented to dig deep into users' specialization areas) and multi-disciplinary bibliographic recommendations (which allow users to elicite resources whose topics are tangentially related to their preferences).

3.2.3. *Fuzzy linguistic RSs based on multi-granular FLM*

Three fuzzy linguistic RSs based on multi-granular FLM have been proposed: [103], [104], and [105].

An academic context in which RSs have been applied satisfactorily is in a Technology Transfer Office (TTO). A TTO is responsible for putting into action and managing the activities which generate knowledge and technical and scientific collaboration. So, the main mission in a TTO is to encourage and help the generation of knowledge and its spread and transfer to the society, with the aim of rapidly meeting society needs and demands. TTO is composed by a team of technicians that are experts in technology transfer and provide information about research resources to the researchers and companies, that is bulletins, projects, calls, notices, events, congresses, courses, and so on. This task requires the selection by the expert of suitable researchers to deliver the information. The large increase of research resources is contributing to that TTO experts not being able to spread the information to the suitable users (both researchers and companies). Then TTO experts are in need of tools to help them. In [103] Carlos Porcel, Antonio G. López-Herrera and E. Herrera-Viedma[103] address this problem by proposing the system called "SIRE2IN", i.e., a fuzzy linguistic content-based RS based on a multi-granular FLM [32]. Different label sets defined to represent the different concepts to be assessed for different users in the filtering activity are used as in [93]. All the linguistic information generated in the system is supported by means of a symbolic fuzzy linguistic approach [28,29].

Assuming the same framework considered in [99], that is an academic digital library, in [104] Carlos Porcel, Juan M. Moreno and Enrique Herrera-Viedma define a fuzzy linguistic hybrid RS based on a different multi-granular FLM [33] which is based on the 2-tuple fuzzy linguistic approach [31]. As in [99] this RS can recommend specialized and complementary resources. Furthermore, this system provides users information on university researchers of related areas with the aim of discovering collaboration possibilities and so, to form multi-disciplinary research groups inside university.

In [104] we assume that the user profiles are provided directly by the own users. In [105] Carlos Porcel and Enrique Herrera-Viedma present a new fuzzy linguistic hybrid RS based on a multi-granular FLM [33] that facilitates the acquisition of the user preferences to characterize the user profiles as in [96]. We assume that users provide their preferences by means of incomplete fuzzy linguistic preference relations and tools to manage incomplete information are used to obtain the user profiles [100,101]. In this way, the acquisition of the user profiles is improved.

3.3. *Web quality evaluation based on fuzzy linguistic modelling*

Quality is a key concept in any IAS, and usually, only a user can determine the actual quality of an information item, i.e., the usefulness, pertinence, appropriateness, completeness or utility of that information item with respect to his/her information needs and/or preferences. For this reason, some web quality evaluation tools incorporate FLM tools in order to facilitate the user participation in the quality evaluation process. In this section, we analyze different web quality evaluation approaches designed using FLM:

- Web quality evaluation models based on Symbolic FLM.
- Web quality evaluation models based on 2-Tuple FLM.

3.3.1. *Web quality evaluation models based on symbolic FLM*

We can find three web quality evaluation proposals based on symbolic FLM in the literature: [79], [80], [106] and [107].

In [80] Enrique Herrera-Viedma and Eduardo Peis present a quality evaluation method of SGML documents based on a symbolic FLM [28,29]. We consider that the elements in a DTD (Document Type Definition) are not equally informative. This is indicated in the DTD by defining linguistic importance attributes to the more meaningful elements of DTD chosen. Then, considering that the evaluation scheme is composed by a finite number of elements of DTD, the quality evaluation method generates linguistic quality assessments from linguistic evaluation judgements provided by different users on those meaningful elements of DTD. To do so, the LWA (Linguistic Weighted Averaging) [28] and LOWA (Linguistic Ordered Weighted Averaging) [29] operators are used in order to obtain quality assessments taking into account the fuzzy majority of the judgements provided by the users.

Enrique Herrera-Viedma et al. present in [79] a quality evaluation methodology of Web sites that store documents. This methodology is qualitative and user-oriented because it generates linguistic recommendations on the information quality of the content-driven Web sites based on users' perceptions. It is composed of two main components, an evaluation scheme to analyze the information quality of Web sites, and a measurement method to generate the linguistic quality assessments. The evaluation scheme is based on both technical criteria related to the Web site structure, and criteria related to the content of information on the Web sites. It is user-driven because the chosen criteria are easily understandable by the users, in such a way that Web visitors can assess them by means of linguistic evaluation judgements. The evaluation scheme is defined following the quality evaluation scheme defined for information systems in [70,71,72,73]. The measurement method is user-centered because it generates linguistic quality assessments of the Web sites based on the site visitors' linguistic evaluation judgements. To combine the linguistic evaluation judgements we introduce two new majority guided linguistic aggregation operators, the MLIOWA (Majority guided Linguistic Induced OWA) and weighted MLIOWA operators, that generate the linguistic quality assessments according to the majority of the evaluation judgements provided by different visitors.

In [106] Enrique Herrera-Viedma et al. introduce a fuzzy linguistic quality evaluation model to measure the quality of Web sites that store XML documents. In this model we combine the evaluation scheme of SGML documents given in [80] together with the evaluation scheme of web sites defined in [79]. This model evaluates the information quality of Web sites using only users perceptions, and therefore it is user centered. Fuzzy linguistic techniques are involved in the quality evaluation process to create a user-friendly framework. This model is composed of two main components, an evaluation scheme to analyze the information quality of Web sites and a computing method of quality ratings of Web sites. The evaluation scheme presents both technical criteria related to the Web site characteristics, and criteria related to the content of XML documents stored in the Web sites. The quality ratings represent the ability of Web sites to meet user requirements. Linguistic quality ratings are obtained by combining linguistic evaluation judgements provided by Web visitors on the different evaluation criteria. As in [80] the computing method is based on two operators for fuzzy computing with words, the LOWA [29] and LWA [28] operators. The later allows to manage relative importance degrees among quality criteria in the evaluation process. This model uses the power of XML Schema language to improve the representation of documents in the Web with semantic characteristics related to their quality and thus it is useful to search quality resources in XML format. Web site quality ratings could be used by Web retrieval systems to help users to find the highest quality XML resources for their information needs. Additionally, this model could be helpful to Web developers to improve the quality of Web sites from a user point of view.

On the other hand, the technological developments on the Web are having a great influence over the developments on others information access instruments as digital libraries. As the development of digital libraries is to satisfy user needs, user satisfaction is essential for the success of a digital library. In [107] Francisco J. Cabrerizo et al. present a quality evaluation model based on a symbolic FLM [28,29] to evaluate the quality of digital libraries. We present a user-oriented evaluation scheme based on the information quality framework [70,71,72,73] composed of eleven quality criteria grouped in four quality dimensions. The quality evaluation of digital libraries is defined using the user perceptions on the quality of digital services provided through their web sites. The computing method of quality evaluations of digital libraries is based on the LOWA [29] and LWA [28] operators.

3.3.2. Web quality evaluation models based on 2-tuple FLM

In this globalized world, the extraordinary importance that the health web sites are taking on patients and physicians as a source of information emphasize all those matters related to the evaluation of quality on the Web. The quality assessment of health-related web sites becomes especially relevant because their use implies the existence of a wide range of threats which can affect people's health. In [108] Juan M. Moreno et al. present a quality evaluation methodology for assessing quality of health-related web sites based on the 2-tuple FLM. To identify the users' perspective quality criteria set, a qualitative research has been carried out using the focus groups technique [109]. As a result of this research, we obtain a user-driven Quality Criteria Framework composed of thirty criteria and grouped in five quality dimensions. The measurement method generates linguistic assessments considering the visitors' judgements with respect to those quality criteria. The combination of the linguistic judgements is implemented without loss of information by applying the 2-tuple Linguistic Weighted Averaging Operators.

4. Conclusions

In this paper, we have shown that the FLM has demonstrated to be a useful tool to improve the performance of different technologies related to the IASs, i.e., IRSs, RSs and Web quality models. The main novelty of the application of FLM in IASs is that it allows users to access to quality and relevant information in a flexible and user-adapted way.

If we analyze the different linguistic approaches of technologies related to IASs we can easily observe that two main research groups have developed them: the Italian group composed by Gloria Bordogna and Gabriella Pasi at the CNR - IDPA and the Spanish group led by Enrique Herrera-Viedma at the University of Granada.

Finally, we should point out that many of these approaches have been published in important journals of SC tools, as *Fuzzy Sets and Systems* [88,102], *Information Sciences* [82], *Int. J. of Approximating Reasoning* [83,93,106], *Int. J. of Intelligent Systems* [91,38], etc. However, it is really important to emphasize that many of them have been published in important journals outside of SC scope and related to the Library and Information Sciences as *J. of American Society of Information Sciences and Technology* [81,85,79], *Information Processing & Management* [80], *Information Research* [99], *Information Retrieval* [89], and *Information Technology and Libraries* [95]. Furthermore, if we do a citation study of some papers according to the ISI Web of Science[§](WoS) and Google Scholar[¶]we can see that they have been well cited and recognized in scientific literature (see Table 1[‖]).

[§]ISI Web of Science. http://scientific.thomson.com/products/wos/.
[¶]http://scholar.google.com
[‖]Checked in 15 May, 2010

Paper	WoS	Google Scholar
Reference [81]	72	182
Reference [85]	50	118
Reference [93]	40	71
Reference [79]	21	49
Reference [83]	18	42

Table 1. Citations of papers.

Finally, we should point out some future works that could be potential applications of FLM in the development of IASs:

1. In IRSs we should study how to apply linguistic weighting tools in Boolean queries to control the behaviour of the aggregation operators used to model the action of the logical connectives AND and OR.
2. In RSs we should study how to apply FLM in the new paradigm of RSs, i.e., trust based RSs.
3. In Web quality evaluation we should extend the application of the FLM to e-commerce activities, in order to consider the user perceptions in the quality evaluation models.

Acknowledgments

This paper has been developed with the financing of FEDER funds in FUZZYLING project (TIN2007-61079), PETRI project (PET2007-0460), project of Ministry of Public Works (90/07) and Excellence Andalusian Project (TIC5299).

References

1. L.A. Zadeh. What is Soft Computing?. *Soft Comp.*,1(1):1, 1997.

2. G. Bordogna and G. Pasi. Modeling Vagueness in Information Retrieval. *Lect. Notes on Inform. Retrieval (LNCS 1980). Springer Verlag*, 196:197, 2000.
3. F. Crestani and G. Pasi. *Soft Computing in Information Retrieval: Techniques and Applications*. Physica Verlag,2000.
4. F. Crestani and G. Pasi. Handling Vagueness, Subjectivity, and Imprecision in Information Access: An Introduction to the Special Issue. *Inf. Process. Manag.*, 39(2):161–165, 2003.
5. E. Herrera-Viedma, F. Crestani, and G. Pasi. *Soft Computing for Web Information Retrieval*. Physica Verlag,2006.
6. E. Herrera-Viedma and G. Pasi. Soft Approaches to Information Retrieval and Information Access on the Web: An Introduction to the Special Topic Section: Special Topic Section on Soft Approaches to Information Retrieval and Information Access on the Web. *J. Am. Soc. Inf. Sci. Tech.*, 57(4):514, 2006.
7. D.H. Kraft, G. Pasi, and G. Bordogna. Vagueness and Uncertainty in Information Retrieval: How Can Fuzzy Sets Help?. In *Proc. of the 2006 Inter. Workshop on Research Issues in Digital Libraries Table of Contents*. ACM New York, NY, USA, 2006.
8. A.G. López-Herrera and E. Herrera-Viedma. Guest Editorial: Soft Computing on the Web. *Soft Comp.*,14(8):783–784, 2010.
9. L.A. Zadeh. The Concept of a Linguistic Variable and Its Applications to Approximate Reasoning. Part I. *Inf. Sci.*,8:199–249, 1975.
10. L.A. Zadeh. The Concept of a Linguistic Variable and Its Applications to Approximate Reasoning.Part II. *Inf. Sci.*,8:301–357, 1975.
11. L.A. Zadeh. The Concept of a Linguistic Variable and Its Applications to Approximate Reasoning.Part III. *Inf. Sci.*,9:43–80, 1975.
12. S. Alonso, F. Chiclana, E. Herrera-Viedma, and F. Herrera. Computing With Words in Decision Making: Foundations, Trends and Prospects. *Fuzzy Optimization and Decision Making*, 8(4):337–364, 2009.
13. B. Arfi. Fuzzy Decision Making in Politics: A Linguistic Fuzzy-Set Approach (LFSA). *Political Analysis*, 13(1):23–56, 2005.
14. A.H. Gharehgozli, M. Rabbani, N. Zaerpour, and J. Razmi. A comprehensive decision-making structure for acceptance/rejection of incoming orders in make-to-order environments. *Inter. J. of Advanced Manufacturing Technology*, 39(9-10):1016–1032, 2008.
15. F. Herrera, E. Herrera-Viedma, and J.L. Verdegay. A Linguistic Decision Process in Group Decision Making. *Group Decision and Negotiation*, 5:165–176, 1996.
16. F. Herrera, E. Herrera-Viedma, and J.L. Verdegay. Choice Processesfor Nonhomogeneous Group Decision Making in Linguistic Setting. *Fuzzy Sets and Sys.*, 94:287–308, 1998.
17. E. Herrera-Viedma. An Information Retrieval Model with Ordinal Linguistic Weighted Queries Based on Two Weighting Elements. *Inter. J. of Uncertainty, Fuzziness and Knowledge-Based Sys.*, 9(4):77–87, 2001.
18. T.-H. Hsu and L.-Z. Lin. Using fuzzy set theoretic techniques to analyze travel risk: An empirical study. *Tourism Manag.*, 27(5):968–981, 2006.
19. Y.-P. Jiang, Z.-P. Fan, and J. Ma. A method for group decision making with multi-granularity linguistic assessment information. *Inf. Sci.*,178(4):1098–1109, 2008.
20. Z. Xu. An Approach Based on the Uncertain LOWG and Induced Uncertain LOWG Operators to Group Decision Making with Uncertain Multiplicative Linguistic Preference Relations. *Decision Support Sys.*, 41(6):488–499, 2006.
21. R. Degani and G. Bortolan. The Problem of Linguistic Approximation in Clinical Decision Making. *Inter. J. Approximate Reasoning*, 2(2):143–162, 1988.
22. R.M. Tong and P.P. Bonissone. A linguistic approach to decisionmaking with fuzzy sets. *IEEE Trans. Sys., Man and Cybernetics*, 10(11):716–723, 1980.
23. R.R. Yager. On there translation process in Zadeh's paradigm of computing with words. *IEEE Trans. Sys., Man, and Cybernetics - Part B: Cybernetics*,34(2):1184–1195, 2004.
24. J.M. Mendel. An architecture for making judgement using computing with words. *Inter. J. of Applied Math. and Computer Sci.*, 12(3):325–335, 2002.
25. J.M. Mendel. Historical reflections on perceptual computing. In *Proc. of the 8th Inter. FLINS Conf. on Computational Intell. in Decision and Control*, pages 181–186, 2008.
26. I.B. Türkşen. Type 2 representation and reasoning for cww. *Fuzzy Sets and Sys*, 127:17–36, 2002.
27. I.B. Türkşen. Meta-linguistic axioms as a foundation for computing with words. *Inf. Sci.*,177(2):332–359, 2007.
28. F. Herrera and E. Herrera-Viedma. AggregationOperators for Linguistic Weighted Information. *IEEE Trans. on Sys., Man, and Cybernetics, Part A: Sys. and Humans*, 27:646–656, 1997.
29. F. Herrera, E. Herrera-Viedma, and J.L. Verdegay. Direct Approach Processes in Group Decision Making Using Linguistic OWA Operators. *Fuzzy Sets and Sys.*, 79:175–190, 1996.
30. R.R. Yager. Concepts, Theory and Techniques, A New Methodology for Ordinal Multi-Objective Decision Based on Fuzzy Sets. *Decision Sci.*, 12:589–600, 1981.
31. F. Herrera and L. Martínez. A 2-tuple Fuzzy Linguistic Representation Model for Computing With Words. *IEEE Trans. on Fuzzy Sys.*, 8(6):746–752, 2000.

32. F. Herrera, E. Herrera-Viedma, and L. Martínez. A Fusion Approach for Managing Multi-Granularity Linguistic Terms Sets in Decision Making. *Fuzzy Sets and Sys.*, 114(1):43–58, 2000.
33. F. Herrera and L. Martínez. A Model Based on Linguistic 2-tuples for Dealing with Multi-granularity Hierarchical Linguistic Contexts in Multiexpert Decision-Making. *IEEE Trans. on Sys., Man and Cybernetics. Part B: Cybernetics*, 31(2):227–234, 2001.
34. L. Martínez, M. Espinilla, and L. G. Perez. A Linguistic Multigranular Sensory Evaluation Model for Olive Oil. *Inter. J. of Computational Intell. Sys.*, 1(2):148–158, 2008.
35. L. Martínez, M. J. Barranco, L. G. Perez, and M. Espinilla. A Knowledge Based Recommender System with Multigranular Linguistic Information. *Inter. J. of Computational Intell. Sys.*, 1(3):225–236, 2008.
36. F.J. Cabrerizo, S. Alonso, and E. Herrera-Viedma. A Consensus Model for Group Decision Making Problems with Unbalanced Fuzzy Linguistic Information. *Inter. J. Inf. Tech. & Decision Making*, 8(1), 2009.
37. F. Herrera, E. Herrera-Viedma, and L. Martínez. A Fuzzy Linguistic Methodology to Deal with Unbalanced Linguistic Term Sets. *IEEE Trans. on Fuzzy Sys.*, 16(2):354–370, 2008.
38. E. Herrera-Viedma and A.G.López-Herrera. A Model of Information Retrieval System with Unbalanced Fuzzy Linguistic Information. *Inter. J. Intell. Sys.*, 22(11):1197–1214, 2007.
39. R. Baeza-Yates and B. Ribeiro-Neto. *Modern Information Retrieval*. Addison-Wesley,1999.
40. G. Salton and M.J. McGill. *An Introduction to Modern Information Retrieval*. McGraw-Hill, 1983.
41. H. Sun. Efficient algorithms for spatial configuration information retrieval. *Knowledge-Based Sys.*, 22(6):403–409, 2009.
42. U. Hanani, B. Shapira, and P. Shoval. Information Filtering: Overview of Issues, Research and Systems. *User Modeling and User-Adapted Interaction*, 11:203–259, 2001.
43. P. Reisnick and H.R. Varian. Recommender Systems. *Communications of the ACM*, 40(3):56–59, 1997.
44. J. Bobadilla, F. Serradilla, A. Hernando, and MovieLens. Collaborative filtering adapted to recommender systems of e-learning. *Knowledge-Based Sys.*, 22(4):261–265, 2009.
45. A.M. Aladwani and P.C. Palvia. Developing and Validating an Instrument for Measuring User Perceived Web Quality. *Inf. & Manag.*, 39:467–476, 2002.
46. L. Olsina and G. Rossi. Measuring Web Application Quality With WebQEM. *IEEE Multimedia*, 2002:20–29, 2002.
47. S.Y. Rieh. Judgment of Information Quality and Cognitive Authority in the Web *J. Amer. Soc. Inf. Sci. Tech.*, 53(2):145–161, 2002.
48. R. Baeza-Yates. Information Retrieval in the Web: Beyond Current Search Engines. *Inter. J. of Approximate Reasoning*, 34(2–3):97–104, 2003.
49. C.J. Van Rijsbergen. *Information Retrieval*. Butterworth, 1979.
50. L.A. Zadeh. Fuzzy Sets. *Inf. and Control*, 8(3):338–353, 1965.
51. S.M. Chen and Wang J.Y. Document retrieval using knowledge-based fuzzy information retrieval techniques. *IEEE Trans. on Sys., Man, and Cybernetics*, 25:793–802, 1995.
52. V. Cross. Fuzzy Information Retrieval. *J. Intell. Inf. Sys.*, 3:29–56, 1994.
53. D.H. Kraft and D.A. Buell. Fuzzy sets and generalized boolean retrieval systems. *Int. J. Man-Machine Studies*, 19:45–56, 1983.
54. D. Lucarella and R. Morara. FIRST: Fuzzy Information Retrieval System. *J. Inf. Sci.*, 17:81–91, 1991.
55. T. Radecki. Fuzzy Set Theorical Approach to Document Retrieval. *Inf. Process. Manag.*, 15:247–260, 1979.
56. V. Tahani. A Fuzzy Model of Document Retrieval Systems. *Inf. Process. Manag.*, 12:177–187, 1976.
57. A. Bookstein. Fuzzy request: An approach to weighted boolean searches. *J. Amer. Soc. Inf. Sci. Tech.*, 31(4):240–247, 1980.
58. G. Bordogna, P. Carrara, and G. Pasi. Query Term Weights as Constraints in Fuzzy Information Retrieval. *Inf. Process. Manag.*, 27(1):15–26, 1991.
59. D.A. Buell and D.H. Kraft. A Model for a Weighted Retrieval System. *J. Amer. Soc. Inf. Sci. Tech.*, 32(3):211–216, 1981.
60. D. Buell and D.H. Kraft. Threshold Values and Boolean Retrieval Systems. *Inf. Process. Manag.*, 17:127–136, 1981.
61. R. Burke. Hybrid web Recommender Systems. *The Adaptive Web*, pages 377–408, 2007.
62. L.M. Quiroga and J. Mostafa. An Experiment in Building Profiles in Information Filtering: the Role of Context of User Relevance Feedback. *Inf. Process. Manag.*, 38(5):671–694, 2002.
63. J. Furner. On Recommending. *J. Amer. Soc. Inf. Sci. Tech.*, 53(9):747–763, 2002.
64. R.R. Yager. Fuzzy Logic Methods in Recommender Systems. *Fuzzy Sets and Sys.*, 136(2):133–149, 2003.
65. A. Popescul, L.H. Ungar, D.M. Pennock, and S. Lawrence. Probabilistic Models for Unified Collaborative and Content-Based Recommendation in Sparse-Data Environments. In *Proc. of the 17th Conf. in Uncertainty in A.I.*, pages 437–444. Citeseer, 2001.
66. P. Perny and J.D. Zucker. Collaborative filtering methods based on fuzzy preference relations. In *Proc. of the EUROFUSE-SIC*, pages 279–285. Kluwer, 1999.
67. S. Lawrence and C. Giles. Searching the Web: Gen-

eral and Scientific Information Access. *IEEE Communications*, 37(1):116–122, 1998.
68. G. Tyburski. Get Smart About Web Site I.Q., 2002. Retrieved April 1,2010, from http://searchenginewatch.com/searchday/article.php/2159621.
69. M. Gertz, M.T. Ozsu, G. Saake, and K.U. Sattler. Report on the Dagstuhl Seminar "Data Quality on the Web". *Sigmod Record*, 31(1):127–132, 2004.
70. K.T. Huang, Y.W. Lee, and R.Y. Wang. *Quality Information and Knowledge*. Prentice Hall PTR Upper Saddle River, NJ, USA, 1998.
71. Y.W. Lee, D.M. Strong, B.K. Kahn, and R.Y. Wang. AIMQ: A Methodology for Information Quality Assessment. *Infor. & Manag.*, 40(2):133–146, 2002.
72. D.M. Strong, Y.W. Lee, and R.Y. Wang. Data Quality in Context. *Communications of the ACM*, 40(5):103–110, 1997.
73. R.Y. Wang and D.M. Strong. Beyond Accuracy: What Data Quality Means to Data Consumers. *J. Manag. Inf. Sys.*, 12(4):5–34, 1996.
74. E.K. Huizingh. The Content and Design of Web Sites: An Empirical Study. *Inf. & Manag.*, 37(3):123–134, 2000.
75. F. Naumann. Quality-Driven Query Answering for Integrated Information Systems. *Lect. Notes in Computer Sci.*, 2261, 2002.
76. S.S. Robbins and A.C. Stylianou. Global Corporate Web Sites: An Empirical Investigation of Content and Design. *Inf. & Manag.*, 40(2):205–212, 2003.
77. M. Bovee, R.J. Srivastava, and B. Mak. A Conceptual Framework and Belief-Function Approach to Assessing Overall Information Quality. *Inter. J. Intell. Sys.*, 18:51–74, 2003.
78. D. Dhyani, W. Keong Ng, and S.S. Bhowmick. A Survey of Web Metrics. *ACM Computing Surveys*, 34(4):469–503, 2002.
79. E. Herrera-Viedma, G. Pasi, A.G. López-Herrera, and C. Porcel. Evaluating the Information Quality of Web Sites: A Qualitative Methodology Based on Fuzzy Computing With Words. *J. Amer. Soc. Inf. Sci. Tech.*, 57(4):538–549, 2006.
80. E. Herrera-Viedma and E. Peis. Evaluating the Informative Quality of Documents in SGML Format from Judgements by Means of Fuzzy Linguistic Techniques Based on Computing With Words. *Inf. Process. Manag.*, 39(2):233–249, 2003.
81. G. Bordogna and G. Pasi. A Fuzzy Linguistic Approach Generalizing Boolean Information Retrieval: A Model and its Evaluation. *J. Amer. Soc. Inf. Sci.*, 44(2):70–82, 1993.
82. D.H. Kraft, G. Bordogna, and G. Pasi Extended Fuzzy Linguistic Approach to Generalize Boolean Information Retrieval. *Inf. Sci.*, 2(3):119–134, 1994.
83. G. Bordogna and G. Pasi. Controlling Retrieval Through a User-Adaptive Representation of Documents. *Inter. J. of Approximate Reasoning*, 12:317–339, 1995.
84. L.A. Zadeh. Computational Approach to Fuzzy Quantifiers in Natural Languages. *Computers & Maths. with Apps.*, 9(1):149–184, 1983.
85. E. Herrera-Viedma. Modeling the Retrieval Process of an Information Retrieval System Using an Ordinal Fuzzy Linguistic Approach. *J. Amer. Soc. Inf. Sci. Tech.*, 52(6):460–475, 2001.
86. G. Bordogna and G. Pasi. An Ordinal Information Retrieval Model. *Inter. J. of Uncertainty, Fuzziness and Knowledge-Based Sys.*, 9:63–75, 2001.
87. C.A. Coello, G.B. Lamont, and D.A. Van Veldhuizen. *Evolutionary Algorithms for Solving Multi-Objective Problems*. Springer-Verlag New York Inc, 2007.
88. A.G. López-Herrera, E. Herrera-Viedma, and F. Herrera. Applying Multi-objective Evolutionary Algorithms to the Automatic Learning of Extended Boolean Queries in Fuzzy Ordinal Linguistic Information Retrieval Systems. *Fuzzy Sets and Sys.*, 160(15):2192–2205, 2009.
89. E. Herrera-Viedma, A.G. López-Herrera, S. Alonso, J.M. Moreno, F.J. Cabrerizo, and C. Porcel. A Computer-Supported Learning System to Help Teachers to Teach Fuzzy Information Retrieval Systems. *Inf. Retrieval*, 12:179–200, 2009.
90. M. Delgado, F. Herrera, E. Herrera-Viedma, M.J. Martin Bautista, L. Martínez, and M.A.Vila. A Communication Model Based on the 2-tuple Fuzzy Linguistic Representation for a Distributed Intelligent Agent System on Internet. *Soft Comp.*,6(5):320–328, 2002.
91. E. Herrera-Viedma, A.G. López-Herrera, and C. Porcel. Tuning the Matching Function for a Threshold Weighting Semantics in a Linguistic Information Retrieval System. *Inter. J. Intell. Sys.*, 20(9):921–937, 2005.
92. E. Herrera-Viedma, A.G. López-Herrera, M. Luque, and C. Porcel. A Fuzzy Linguistic IRS Model Based on a 2-Tuple Fuzzy Linguistic Approach. *Inter. J. of Uncertainty, Fuzziness and Knowledge-based Sys.*, 15(2):225–250, 2007.
93. E. Herrera-Viedma, O. Cordón, M. Luque, A.G. López, and A.M. Muñoz. A Model of Fuzzy Linguistic IRS Based on Multi-Granular Linguistic Information. *Inter. J. of Approximate Reasoning*, 34:221–239, 2003.
94. Z. Pei, D. Ruan, Y. Xu, and J. Liu. Handling Linguistic Web Information Based on a Multi-Agent System. *Inter. J. Intell. Sys.*, 22(5):435–454, 2007.
95. J.M. Morales del Castillo, R. Pedraza-Jiménez, A.A. Ruíz-Rodríguez, E. Peis, and E. Herrera-Viedma. A Semantic Model of Selective Dissemination of Information for Digital Libraries. *Inf. Tech.and Libraries*, 28(1):22–31, 2009.

96. J.M. Morales del Castillo, E. Peis, A.A. Ruíz-Rodríguez, and E. Herrera-Viedma. A Web-Based Service for the Elicitation of Resources in the Biomedical Domain. *Inter. J. of Intell. Sys.*, In Press 2010.
97. A.F. Smeaton and J. Callan. Personalisation and Recommender Systems in Digital Libraries. *Inter. J. Digital Libraries*, 5:299–308, 2005.
98. Berners-Lee, T. and Hendler, J. and Lassila, O. The semantic web: a new form of web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific Amer.*, 279(5), 2001.
99. J.M. Morales del Castillo, E. Peis, J.M. Moreno, and E. Herrera-Viedma. D-Fussion: A Semantic Selective Disssemination of Information Service for the Research Community in Digital Libraries. *Inf. Research*, 14(2):paper 398, 2009.
100. S. Alonso, F. Chiclana, F. Herrera, E. Herrera-Viedma, J. Alcalá, and C. Porcel. A Consistency Based Procedure to Estimate Missing Pairwise Preference Values. *Inter. J. Intell. Sys.*, 23:155–175, 2008.
101. S. Alonso, F.J. Cabrerizo, F. Chiclana, F. Herrera, and E. Herrera-Viedma. Group Decision-Making with Incomplete Fuzzy Linguistic Preference Relations. *Inter. J. Intell. Sys.*, 24(2):201–222, 2009.
102. E. Herrera-Viedma, F. Herrera, L. Martínez, J.C. Herrera, and A.G. López. Incorporating Filtering Techniques in a Fuzzy Linguistic Multi-Agent Model for Information Gathering on the Web. *Fuzzy Sets and Sys.*, 148(1):61–83, 2004.
103. C. Porcel, A.G. López-Herrera, and E. Herrera-Viedma. A Recommender System for Research Resources Based on Fuzzy Linguistic Modeling. *Expert Sys. with Apps.*, 36(3):5173–5183, 2009.
104. C. Porcel, J.M. Moreno, and E. Herrera-Viedma. A Multi-Disciplinar Recommender System to Advice Research Resources in University Digital Libraries. *Expert Sys. with Apps.*, 36(10):12520–12528, 2009.
105. C. Porcel and E. Herrera-Viedma. Dealing with Incomplete Information in a Fuzzy Linguistic Recommender System to Disseminate Information in University Digital Libraries. *Knowledge-Based Sys.*, 23:32–39, 2010.
106. E. Herrera-Viedma, E. Peis, J.M. Morales del Castillo, S. Alonso, and K. Anaya. A Fuzzy Linguistic Model to Evaluate the Quality of Web Sites that Store XML Documents. *Inter. J. Approximate Reasoning*, 46(1):226–253, 2007.
107. F.J. Cabrerizo, J. López-Gijón, A.A. Ruíz-Rodríguez, and E. Herrera-Viedma. A Model Based on Fuzzy Linguistic Information to Evaluate the Quality of Digital Libraries. *Inter. J. Inf. Tech. & Decision Making*, 9(3):455–472, 2010.
108. J.M. Moreno, J.M. Morales del Castillo, C. Porcel, and E. Herrera-Viedma. A Quality Evaluation Methodology for Health-Related Websites Based on a 2-tuple Fuzzy Linguistic Approach. *Soft Comp.*, 14(8):887-897, 2010.
109. D.W. Stewart and P.N.Shamdasani. *Focus Groups: Theory and Practice (Applied Social Research Methods)*. Sage Publications,1990.