

SOLAR: Social Link Advanced Recommendation System

Ángel García-Crespo^a, Ricardo Colomo-Palacios^{a,*}, Juan Miguel Gómez-Berbís^a,
Francisco García-Sánchez^b

^a Escuela Politécnica Superior, Universidad Carlos III de Madrid, Avda. de la Universidad 30, 28911 Leganés, Madrid, Spain

^b Departamento de Informática y Sistemas, Facultad de Informática, Universidad de Murcia, 30071 Espinardo (Murcia), Spain



ARTICLE INFO

ABSTRACT

Keywords:
Hyperlinks
Semantic links
Semantic Web
Affect Grid
Social Web

In today's information society, precise descriptions of the massive volume of online content available are crucial for responding to user needs adequately and efficiently. The Semantic Web Paradigm has recently advanced across many domains for the assignment of metadata to Internet content, in order to define it with explicit, machine-readable meaning. This content has become so extensive that it must be refined according to user preferences to avoid information overload. The current paper proposes a framework for the association of semantic data to webpage links based on a specific domain ontology, additionally permitting the user to express his opinion regarding his emotions about the content of the link. This data is further exploited to suggest additional links to the user, based on the semantic metadata and the level of user satisfaction with previously viewed content. A comprehensive evaluation of the tool has demonstrated a high level of user satisfaction with the features of the system.

1. Introduction

Since its humble beginnings, the Internet has gained vast importance in today's society, both in terms of consumer reach and the volume of fundamental information it contains for millions of users worldwide. Transforming from a host-to-host network connection in ARPANET [1] on October 29, 1969, it has advanced to become what Tim Berners-Lee has termed the "Giant Global Graph" [2]. Over the years, connections were made between the data in this graph through the use of hyperlinks. This data is constantly increasing, concurrently increasing the number of hyperlinks which connect the data. However, in order to efficiently access this data, nowadays users require information filters and new ways to access the information.

The evolution of the Web from Web 1.0 to Web 2.0, and to what has recently been termed by some as Web 3.0, has caused a paradigm shift in a user's access to and control of websites. The Web 2.0 phenomenon made the Web social, initiating an explosion in the number of users of the Web, thus empowering them with a huge autonomy in adding content to webpages, labeling the content, creating folksonomies of tags, and finally, leading to millions of users constructing their own webpages [3].

Logically, the result of this movement was a significant increase in the number of webpages available. Surfing through endless links is no longer an efficient method for finding relevant information. This not only applies to the standard user, all owners of webpages, whether organizations or consumers, lose target audiences as users of the Web are unable to access their pages in an efficient time frame. From the perspective of the users of webpages, both ordinary users and businesses also operate inefficiently, losing time when searching for relevant pages. Thus, a new approach is required which identifies links defined using specific data. The focus of the current paper is a platform which aids the user to encounter relevant hyperlinks. This work proposes the use of a combination of Semantic Web technology to assign meaning to the links and generate directly related links, and the application of an established theory of emotion levels to suggest additional links from the same subject category to the user, according to his interest. The recommended links are based on the affect valence which the user has previously expressed regarding similar subject matter. In the current section, the paper provides an overview of hyperlinks and an introduction to the Semantic Web concept, which will be presented in further detail subsequently in the paper.

Traditionally, ordinary webpages base the relationship between pages in explicit hyperlinks which lack any type of classification metadata (although some initiatives recently been made; see below for further discussion). The classification of objects, for example textual resources, with semantic metadata or opinion categories has been exploited in other domains, such as the machine learning domain, in particular for text classification using

* Corresponding author.

E-mail addresses: angel.garcia@uc3m.es (Á. García-Crespo),
ricardo.colomo@uc3m.es (R. Colomo-Palacios), juanmiguel.gomez@uc3m.es
(J.M. Gómez-Berbís), frgarcia@um.es (F. García-Sánchez).

Natural Language Processing Techniques. However, Semantic Web and sentiment analysis (or opinion mining) technologies have not been exploited until now to help the user with a crucial activity for efficient Internet use, that is, to select only the links relevant to his/her requirements. If the relation between Internet resources were represented by means of semantics, the application would be able to provide mechanisms to semantically navigate between related resources with real meaning. Adding sentiment to links additionally enables the application to provide the user with links which may be of interest to him/her. As ordinary hyperlinks are not sufficient to display semantic data to users, another approach is required to offer the user all the information semantically when viewing data. The user interface must enable navigation to semantically related items. In order to achieve this, semantic links are proposed, which are ordinary hyperlinks in appearance but built upon semantic information. This semantic information, consisting both of an ontological concept to which a particular part of the content of the link refers, will lead the user to pages with semantically similar content by means of recommendations.

The current work proposes to annotate links within social content such as blogs with concepts from a specific domain ontology and a measure of the well-known Affect Grid psychosocial theory proposed by Russell, Weiss and Mendelsohn [4].

The section which follows provides an overview of the State of the Art of hypermedia links and the Semantic Web concept. Section 3 outlines the SOLAR approach, and subsequently a use case is presented to demonstrate the application of the framework in a real scenario. Section 4 analyzes the results of the evaluation of the tool, and Section 5 describes the conclusions and future work.

2. State of the art

As stated by [5], one of the original success factors of the Web is that it provides simple access to the information contained in it. Part of this success was initiated by the use of hyperlinks. Independently of the origin of the concept of hyperlinks, which has still not been concretely determined (see [6] for details), their usefulness for the Internet is highly evident. It is precisely this usefulness and its enormous potential which has been a subject of study for researchers all around the world.

Hypermedia is a concept that encourages authors to structure information as an associative network of nodes and interrelating links [5]. Based on some of the studies regarding the addition of comments to links in hypermedia elements [7], in the 1990s decade some authors (see [8–10]) already proposed the addition of semantics to hypermedia links. At the beginning of this decade, Frei & Schauble [11] stated that the objective of semantic links is to point to similar, more detailed, or additional information on a specific topic. Thus, according to [10] a distinction can be made between semantic links and referential links, whose aim is to make the reading of the document comfortable. The advent of the Semantic Web has evoked substantial changes in the field of searching and referencing information. The term “Semantic Web” was coined by Berners-Lee, Hendler and Lassila [12] to describe the evolution from a document based web towards a new paradigm that includes data and information for computers to manipulate. Ontologies [13] are the technological cornerstones of the Semantic Web, because they provide structured vocabularies that describe a formal specification of a shared conceptualization. The fundamental aim of the Semantic Web is to answer the ever-growing need for data integration on the Web. The benefit of adding semantics consists of bridging nomenclature and terminological inconsistencies to include underlying meanings in a unified manner. Given that a universally shared data format is not likely to arise and diffuse, the Semantic Web provides an alternative solution to represent the comprehensive meaning of integrated information and promises

to lead to efficient data managing by establishing a common understanding [14].

A Semantic Link Network (SLN) [15] is an extension of the hyperlink network formed by attaching semantic factors to links. The SLN was proposed as a semantic data model for organizing various Web resources by extending the Web's hyperlink to a semantic link [16]. The semantic factors are concepts reflecting relations which are commonly searched for by users. Compared with other approaches to the Semantic Web, SLN can better inherit the characteristics of the World Wide Web [15] and raise the efficiency of Peer to Peer (P2P) networks [17]. According to [18], a SLN is a directed network, where the nodes are content and the edges are the typed semantic links. The principal structure of the SLN is a sequential chain that connects the main fragments of the document from the beginning to the end node. The use of SLN in various types of applications has demonstrated notably satisfactory results. Applications of this technology to academic citation are discussed in [18–20,16,21]. Apart from the works by Zhuge, Pons has successfully applied the SLN to object prefetching and achieved a better result than other approaches [22].

On the other hand, studies of emotion and affect in human beings have an established history which originates from philosophy. As a result of this tradition, and using their own work as a basis [23], Russell, Weiss and Mendelsohn [4] proposed a measure of affect which had a profound impact in the field of social psychology. They termed the measure the Affect Grid, a scale designed as a quick means of assessing affect along the dimensions of pleasure–displeasure and arousal–sleepiness on a 1–9 scale. The Affect Grid may prove to be the instrument of choice when subjects are called on to make affective judgments in rapid succession or to make a large number of judgments, especially when those judgments are to be aggregated [24]. Using this scale, researchers can collect emotion ratings from stress and tension to calm, relaxation or serenity, and from ecstasy, excitement and joy to depression, melancholy, sadness, and gloom.

According to the studies of these authors, The Affect Grid is potentially suitable for any study that requires judgments about affect of either a descriptive or a subjective kind. Taking into account that the objective of the current work is precisely to assign a numerical measure to links such that a value of their affect can be obtained, the application of the tool was considered ideal for the objectives of the current work.

A number of previous studies exist regarding the use of emotional valence or affect in the Semantic Web domain. In [24], it is proposed to quantify the links in a semantic network with a Boolean value of positive or negative sentiment. As previously mentioned, a different approach is the derivation of sentiment using Natural Language Processing Techniques, where more extensive emotional descriptions may be obtained. [25] proposes the classification of the emotions of customer feedback on the basis of an ontology of consumer emotions. Other works present further initiatives, such as the analysis of emotion in texts [26], or speech recognition [27].

Due to the vast amount of information available Internet users need of mechanisms to manage information. As a result of this, from decades ago uses rate internet content (e.g. products, buyers, sellers, pages, answers ...). This kind of feedback, in the form of summative or formative assessment can be employed by other users to support decision process or, just to surf the net. However, adding semantics to this feedback can be seen as a newer trend. Thus, in the Semantic Web world, there are movements to enrich what people are saying by adding semantics to their links either manually or automatically. In this area the annotation of semantic wikis can be seen as a valid antecedent [28–30], but, compared with mentioned initiatives, the SOLAR approach takes into account emotions of the link-makers to annotate links and enrich them taking full advantage of the gradient of emotion.

The aim of SOLAR is to rank the emotions which the links evoke in the user in a simple and standardized form. Additionally, the use of semantics is a further advantage for the creation of networks of semantically defined links, and emotionally annotated in the field of the contents generated on the social Web.

3. The SOLAR approach

This section presents the SOLAR approach, a Social Link Advanced Recommendation System based on the concept of SLN, which fosters the potential of linking related data on the Web. Fundamentally, data discovery on the Web benefits from the “linked data” approach of referencing resource Uniform Resource Identifier (URI)s, but its potential and efficiency in terms of knowledge management is leveraged by the use of a number of “semantic links” that show non-explicit relationships among concepts through well-defined and well-structured Resource Description Framework (RDF) [31] based relationships.

Ten years after becoming a World Wide Web Consortium (W3C) recommendation, RDF is used to represent data in a number of applications, including business environments such as the Oracle Application Server [32], which turns corporate database generated data into RDF triples. Hence, given that RDF has become a widely accepted formalism for representing semantic relationships and its XML (Extensible Markup Language) [33] syntax is very flexible and scalable, in SOLAR it is deemed an efficient formalism to represent SLN [15] in the context of its architecture.

Obviously, there is a whole lattice of frameworks associating Semantic Data to webpage links. However, the SOLAR approach relies on a two-pronged value-added strategy. On the one hand, apart from simple associations, SOLAR derives, infers or suggests a set of new “hidden relationships”, that is, associations that are not immediately deduced. On the other hand, we also relate emotions and categories to the whole picture, providing a much more detailed and fruitful metadata where the architecture depicted in this section can build on.

The current section details the SOLAR architecture, introducing a novel framework for a knowledge based software infrastructure that implements the SLN concept and demonstrates its advantages. Subsequently, the proof-of-concept implementation is shown and finally, the paper outlines a use case based on the implementation.

3.1. Design principles and architecture

The Social Link Advanced Recommendation System is based on the following three cornerstone principles:

- (1) Following the SLN concept [15,17], structured or semi-structured data can be seen as a data graph and knowledge can be extracted through validating entailment between a graph representing a query and the data graph. In the current work, this model has been termed *hidden relationships*. This concept encompasses finding value-added relationships by means of inference techniques.
- (2) Most relationships can be categorized. A category might belong to taxonomies, a superset of categories where knowledge can also be inferred.
- (3) Valences are human-oriented subjective information values related to feelings. They are a very valuable source of data, since the human being continuously expresses his emotions related to a particular situation or aspect. In SOLAR, a Valence ontology is included which provides subjective information relevant for classifying the information and extracting knowledge using the Affect Grid described in [24].

Following these design principles, the SOLAR architecture is depicted in Fig. 1.

The architecture is partitioned into a three layer perspective. The first layer is user-oriented and related to the front end of the interaction. It splits into the following components:

- Annotation plug-in: Most structured or semi-structured Web resources, including Web 2.0 sites, such as blogs, multimedia sharing or collaborative applications include the possibility of adding a plug-in to their functionalities. In the Web Front layer, an Annotation Plug-in is applied to provide the necessary metadata for the SLN. This plug-in can be implemented as an extension of the dynamic site in environments such as WordPress [34] or Drupal [35] and allow the user annotation of the site with metadata related to Valences (using concepts of the Valence Ontology Repository) and Categories (using concepts of the Category Ontology Repository).
- Search plug-in: This component enables a seamless search through the metadata of the SOLAR architecture. Again, this plug-in could be implemented in WordPress or Drupal and provides Search functionality based on both the Query Engine and the Reasoning engine of the second layer of the SOLAR architecture.
- Navigation interface: As a result of previous research in Digital Libraries [36], a number of RDF based faceted navigation mechanisms were developed that help the user to navigate through the SLN. In summary, the user can browse the resource and navigate through the valences, categories and hidden relationships or links proposed by the SOLAR application.

The second layer provides an Application Logic Manager layer which is divided into the following components:

- Reasoning engine: This component derives facts from a knowledge base, reasoning about the information with the ultimate purpose of formulating new conclusions. In the SOLAR architecture, it consists of an Web Ontology Language (OWL) [37] Description Logics based Reasoner, such as the Renamed ABox and Concept Expression Reasoner (RACER). It uses subsumption to find sets and subsets of annotations based on logical constraints.
- Query engine: The Query Engine component uses the SPARQL (recursive acronym of SPARQL Protocol and RDF Query Language) RDF [38] to make queries into the storage systems of the back end layer. The semantics of the query are defined not by a precise rendering of a formal syntax, but by an interpretation of the most suitable results of the query, given that SOLAR stores mostly RDF triples or OWL DL ontologies which also present an RDF syntax.
- Hidden relationships manager: Given a querying dataset DT with query K, a number of interesting relationships can be logically entailed. Approximations to the semantic meaning of the query K_1 could yield a set of K results, such as $\{K_1 \dots K_N\}$ and a subset of DT can also be inferred. Improving query results by means of more focused and drilled-down queries might return a superset of the original DT where queries are more adjusted. In this component, reasoning and inference based techniques based on these strategies are used to determine added-value relationships. The advantages of this approach will be shown in the Use Case section.

The third layer, the Knowledge Storage, stores the Valence, Category and Domain Ontologies in their Repositories. The distinction among the Repositories is more conceptual than architectural from a software viewpoint. This three repositories are semantic data store systems that enable ontology persistence, querying performed by the Application Logic Manager layer components and offers a higher abstraction layer to enable fast

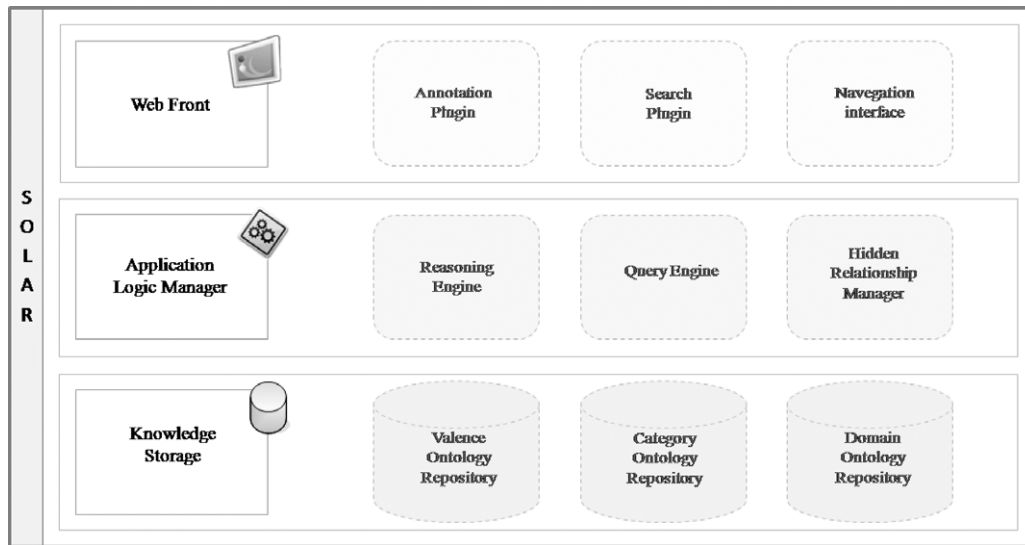


Fig. 1. SOLAR architecture.

storage and retrieval of large amounts of OWL DL ontologies, together with their RDF syntax while keeping a small footprint and a lightweight architecture approach. An example of these systems could be the OpenRDF Sesame RDF Storage system [39], which deal with data and legacy integration.

3.2. Implementation

The SOLAR proof-of-concept implementation originated during the software development phase of the project, based on the architecture guidelines and showed a major concern in a couple of implementation principles. On the one hand, it should be based on open-source widely used platforms and technologies, such as the Wordpress framework (a PHP based blog platform which has gained momentum over the last years becoming the most used according to the blogging community) or the AJAX technology for the Web front.

On the other hand, since Semantic Technologies have reached a certain maturity, the SOLAR implementation benefits from three consolidated ones. Firstly, the Web Ontology Language (OWL) as the cornerstone conceptual and underlying language, particularly its OWL DL flavor, since it builds on the formal foundations of Description Logics (DL), a set of languages whose reasoning can be decidable. As described in Section 3.1 this supports the use of subsumption as the selected reasoning. Secondly, we used the Java based Renamed ABox and Concept Expression Reasoner (RACER) RACER reasoning engine and the Jena framework as the backbone technology for the Application Logic Manager layer, the former for the reasoning and the latter for the RDF Management and SPARQL Querying.

Finally, we worked on a set of PHP based extensions of the plug-ins for annotation in Wordpress, since we realized it would be easy to couple a set of necessary functionalities for the SOLAR implementation. Those functionalities were easily hardcoded and integrated in the plug-ins taking advantage of the flexible and dynamic properties of the Wordpress framework. In addition, we developed an Asynchronous JavaScript and XML (AJAX, for short) based Web interface both in the Search and Navigation software components, since it provided a set of loosely-coupled features to optimize the application efficiency and ease of use from the user standpoint.

3.3. Use case

The solution which SOLAR proposes is based on the use of semantic annotations in hyperlinks which can be added to current communication media, for example, the contents generated within social networks such as blogs. The system exhibits its usefulness in three use case scenarios which are described below. In order to clearly illustrate the use case, it will be detailed using a real social network expressly created by students in the final year of the Computer Science degree in the University Carlos III of Madrid, a blog which was developed to support the student's degree subjects.

Scenario 1: This scenario demonstrates how the annotation of the link is achieved. The user responsible for managing the content of the blog generates a new article or post which contains a link to any content. This is where SOLAR takes place. In this case, the user is adding an annotation to a link related to the keyword PMBOK (Project Management Body of Knowledge), of the current post. This user, who is currently creating the link, after providing the URL, should indicate the following parameters:

- Affect valence:
 - Pleasure: Using a slider, indicate on a scale of 1 to 9 indicate how pleasurable he considers the content of the link which he is currently annotating.
 - Arousal: Indicate on a scale of 1–9 how much excitement the user receives subsequent to viewing the content of the link he is annotating. In this case, the closer the value is to zero, the lower the user rating, and the higher the value is to 10, the higher the match of the link is with user preferences.
- Category: Indicates the category to which the link belongs, within the possibilities which are offered by faceted navigation.

In the context of Use Case 1, the possible categories for the annotation of the link are provided by the ontology selected for the blog. This domain ontology may vary, depending on the theme of the blog; in this case, it is about project management.

Scenario 2: Scenario 2 specifically demonstrates a search by a user for content within the blog using the blog's search functionality. In this case, the objective is to display to the user the links generated according to the search criteria which he has introduced, and, according to the category to which they belong, verify whether they semantically match the links added in the SOLAR system. There are two possible search results:

- Results with a direct semantic connection: Those results which according to the search pattern have a direct semantic relation with the results generated.
- Inferred results: The results obtained based on the possible semantic relations between the search criteria and the links available, on the one hand, and the preference of the user carrying out the search on the other hand.

It is important to mention that the inferred results are only displayed to the user carrying out the search if (s)he is logged to the system, given that these results depend on the user preferences, and these preferences may only be obtained and accessed if the user is logged into the system. If the user does not log into the system, only results with a direct semantic connection are shown, which do not rely on user preferences to demonstrate the results.

Scenario 3: In this scenario, it is assumed that the user is reading any of the posts published in a blog. In the current case, the use case is detailed for the keyword PMBOK in the body of a post. On the one hand, the system generates the results which have a direct semantic relation:

- PMBOK Guide – Fourth Edition.
- ISO 10006.
- PRINCE2.
- IEEE Std 1490-2003.

Moreover, an additional result is displayed, which has been obtained as an inferred result (it is not project management itself but according to user preferences it is shown), as described in Scenario 2:

- Guide to SWEBOK

Thus, five links are provided in the body of the post. Four as “direct” links (have direct semantic relation) and one more, corresponding to SWEBOK, as an inferred link using different style.

4. Evaluation

The subsequent section describes the empirical evaluation of SOLAR, carried out using the application of established statistical tests.

4.1. Research design

Evaluation of the proposal was required in order to determine the level of acceptance of the tool developed. With the objective of calculating the grade of adjustment of the proposal, a study was designed to test the system in relation to different features of the platform, its use and functionalities. The subjects of the study were students in the final year of the Computer Science degree in the University Carlos III of Madrid, who were asked to perform a series of tests to evaluate the capacities of the tool. The students realized a number of tasks relating to the Blogs of one of the subjects within the Computer Science degree, which they themselves maintained. The majority of the content of the blogs was textual; however, they also contained some photo and video files. In this scenario, the students were requested to intersect, that is, connect, annotated links using SOLAR, about which they were subsequently requested to complete a questionnaire specially designed for evaluation purposes.

Once the students had completed the evaluation of the tool as part of the Computer Science subject, they were asked to respond to a questionnaire in which they were solicited for their opinion about the tool in relation to different aspects. Firstly, they were asked about their user experience of SOLAR. Secondly, the user provided his opinion about the performance of the framework. The responses to these questions were coded using a Likert scale ranging from 1 to 4 points, with the following values. 1: Limited,

Table 1

Evaluation of statistical results.

	Average	Std. deviation
User experience	3.06	0.57
Performance	3.00	0.51
Search	2.90	0.79
Inferred search	2.57	0.99
Suggested links	2.86	0.73

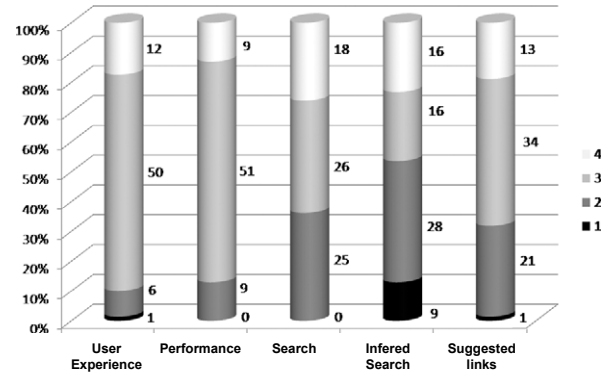


Fig. 2. Score frequencies.

2: Regular, 3: Good, 4: Very Good. Additionally, the users were asked to carry out three specific annotations of content using an ontology representing a concrete domain. Subsequently, the user was asked to perform two searches. The first type of search was a faceted search of the concept which (s)he had introduced, and the second type of search was to perform a query for a related term, to test the inference capacities of the relations in SOLAR. Upon examining the results of the search, the user could quantify the result using a Likert scale with the following values: 1: Unsolved, 2: Low Matching, 3: Near Matching, 4: Right Matching. Lastly, the students were requested to surf the page on which the link had been introduced, with the objective of evaluating the related terms which had been suggested based on the annotation.

4.2. Sample

As mentioned above, the sample was composed of students in the final year of the Computer Science degree of the University Carlos III. In relation to demographic characteristics, the group of subjects consisted of 18 women (26%) and 51 men (74%), with an average age of 24.2. During the administration of the questionnaire, the subjects were aided by a research assistant with the objective that the coding of the questionnaire was carried out accurately, eliminating any student doubts or errors.

4.3. Results

The results of the surveys, which were filled in using printed copies, were subsequently coded in the statistical analysis tool Statistical Package for the Social Sciences (SPSS). In Table 1, the average and standard deviation of the responses offered by the subjects are shown in relation to the questionnaire applied, and the two groups of questions formulated.

The frequency of the scores in the scale defined for the different factors may be viewed in Fig. 2.

Additionally, the research aimed to establish the significant differences between the elements “Search”, “Inferred Search” and “Suggested links”. Taking into account that these three functionalities are characteristic of the search realized by the components of SOLAR, it was additionally aimed to determine the aspects in which the tool could be improved. To establish whether the scores of the

three categories presented significant differences, the statistical method Student's *t*-test (comparison of two means) was used to carry out one-way between-groups analysis of variance (ANOVA) using the tool SPSS. The level of statistical significance was set at 0.05. Three comparisons were made, which generated the following results. The comparison between "Search" and "Inferred Search" presented significant differences ($t(68) = 2.38, p < .05$), as well as the comparison between "Inferred Search" and "Suggested links" ($t(68) = 2.09, p < .05$). On the other hand, the comparison between "Search" and "Suggested links" did not present statistically significant differences ($t(68) = .62, p > .05$). The analysis of the statistical tests indicates that both "Search" and "Suggested Links" presented highly similar values within the evaluation, while "Inferred Search" displayed significant differences with both "Search" and "Suggested Links".

4.4. Discussion

From the perspective of the results, the implementation of SOLAR may be considered a success. Examining the results expressed by the users in the evaluation of the element "User Experience", the elements developed which relate to the graphical perspective hold a very high value for the user. This value presents an average of 3.06 points over 4, with a highly adjusted standard deviation of 0.57, which indicates that, apart from the scoring being high, the agreement between the subjects is acceptable.

"Performance" was judged with equally high values. It presents an average comparable to that of the "User Experience" factor, and a least standard deviation of the factors analyzed of 0.51. It thus emerged that the inclusion of the component developed using a web client can be considered a correct decision.

In relation to the search elements, both faceted and inferred, and the inclusion of suggestions on the Web have displayed satisfactory results. The values presented for "Inferred Search" are slightly inferior to those displayed by "Search" and "Suggested Links". This result is due to the fact that recommendations are not constructed based on direct semantic links, but links inferred drawn from user preferences. On some occasions, the suggestions are not easily processed by the user, as the derivation of user preferences is not trivial. A more refined functionality of "Inferred Search" has been considered as a future research line. Lastly, the values demonstrated by "Search" and "Suggested Links" are considered extremely satisfactory.

5. Conclusions and future work

This paper has presented SOLAR, a Social Link Advanced Recommendation System based on a knowledge based architecture built on three design principles, supporting the Semantic Link Network concept [15,17]. Subsequent to an extensive study of the state of the art, it was determined by the researchers that the combination of Semantic Web technology with emotional valence values could provide added-value features for the Internet user. A real-life case study accompanied by evaluation results verified that users demonstrated satisfaction with the system according to a number of evaluation parameters. The work is supported by a proof-of-concept implementation. Areas for future research include optimizing the data infrastructure, reasoning strategies and including a data crawler. A crawler is a web spider that would also extract information from non-structured data sources on the Web which could extend the scope of SOLAR beyond structured or semi-structured data resources. A latter effort the authors are intending to approach in the future is the inclusion of formative assessment apart from the summative assessment supported now on SOLAR.

This new feature will lead to the inclusion of Natural Language Processing Techniques in order to identify emotions in text and classify them using defined ontologies.

Fundamentally, the most innovative aspect of SOLAR is its philosophy. Interweaving the Semantic Link Network (SLN) concept with current sources of structured data and providing a semantic infrastructure to relate the vast amount of data critical sources on the Web.

Acknowledgements

This work is supported by the Spanish Ministry of Industry, Tourism, and Commerce under the project SONAR (TSI-340000-2007-212), GODO2 (TSI-020100-2008-564) and SONAR2 (TSI-020100-2008-665), under the PIBES project of the Spanish Committee of Education & Science (TEC2006-12365-C02-01) and the MID-CBR project of the Spanish Committee of Education & Science (TIN2006-15140-C03-02).

References

- [1] J. Abbate, *Inventing the Internet*, MIT Press, Cambridge, 1999, pp. 36–111.
- [2] T. Berners-Lee, Giant global graph, <http://dig.csail.mit.edu/breadcrumbs/node/215>, November 2007. Last Verified: June 29th, 2009.
- [3] J.G. Breslin, S. Decker, The future of social networks on the internet: The need for semantics, *IEEE Internet Computing* 11 (6) (2007) 86–90.
- [4] J.A. Russell, A. Weiss, G.A. Mendelsohn, Affect Grid: A single-item scale of pleasure and arousal, *Journal of Personality and Social Psychology* 57 (3) (1989) 493–502.
- [5] M. Bieber, F. Vitali, H. Ashman, V. Balasubramanian, H. Oinas-Kukkonen, Fourth generation hypermedia: Some missing links for the World Wide Web, *International Journal on Human Computer Studies* 47 (1) (1997) 31–65.
- [6] R.H. Stern, Who invented hyperlinks? *IEEE Micro* 21 (1) (2001) 8–10.
- [7] J. Conklin, Hypertext: An introduction and survey, *IEEE Computer* 20 (9) (1987) 17–41.
- [8] H.R. Moskowitz, H. Oinas-Kukkonen, Towards greater flexibility in software design systems through hypermedia functionality, *Information and Software Technology* 39 (6) (1997) 391–397.
- [9] M. Thüring, J. Hannemann, J.M. Haake, Designing for comprehension: A cognitive approach to hypermedia development, *Communications of the ACM* 38 (8) (1995) 57–66.
- [10] H.P. Frei, D. Stieger, The use of semantic links in hypertext information retrieval, *Information Processing & Management* 31 (1) (1995) 1–13.
- [11] H.P. Frei, P. Schläuble, Determining the effectiveness of retrieval algorithms, *Information Processing & Management* 27 (2/3) (1991) 153–164.
- [12] T. Berners-Lee, J. Hendler, O. Lassila, The semantic web, *Scientific American* 284 (5) (2001) 35–40.
- [13] D. Fensel, *Ontologies: A Silver Bullet for Knowledge Management and Electronic Commerce*, Springer, Berlin, Heidelberg, 2002.
- [14] N. Shadbolt, W. Hall, T. Berners-Lee, The semantic web revisited, *IEEE Intelligent Systems* 21 (3) (2006) 96–101.
- [15] H. Zhuge, *The Knowledge Grid*, World Scientific Publishing Co, Singapore, 2004.
- [16] H. Zhuge, Communities and emerging semantics in semantic link network: Discovery and learning, *IEEE Transactions on Knowledge and Data Engineering* 21 (6) (2009) 785–799.
- [17] H. Zhuge, X. Li, Peer-to-peer in metric space and semantic space, *IEEE Transactions on Knowledge and Data Engineering* 19 (6) (2007) 759–771.
- [18] H. Zhuge, Active e-document framework ADF: Model and tool, *Information & Management* 41 (1) (2003) 87–97.
- [19] H. Zhuge, Discovery of knowledge flow in science, *Communications of the ACM* 49 (5) (2006) 101–107.
- [20] H. Zhuge, Retrieve images by understanding semantic links and clustering image fragments, *The Journal of Systems and Software* 73 (3) (2004) 455–466.
- [21] H. Zhuge, Semantics, resource and grid, *Future Generation Computer Systems* 20 (1) (2004) 1–5.
- [22] A.P. Pons, Object prefetching using semantic links, *ACM SIGMIS Database* 37 (1) (2006) 97–109.
- [23] J.A. Russell, A circumplex model of affect, *Journal of Personality and Social Psychology* 39 (1980) 1161–1178.
- [24] Z. Huang, Y. Qiu, Construction and aggregation of citation semantic link network, in: *Proceedings of the 2008 Fourth International Conference on Semantics, Knowledge and Grid*, 2008, pp. 247–254.
- [25] A. García-Crespo, R. Colomo-Palacios, M. Mencke, J.M. Gómez-Berbis, CUSENT: Social sentiment analysis using semantics for customer feedback, in: M.D. Lytras, P. Ordoñez de Pablos (Eds.), *Social Web Evolution: Integrating Semantic Applications and Web 2.0 Technologies*, IGI Global, Hershey, PA, 2009.
- [26] Y. Mathieu, Annotation of emotions and feelings in texts, in: J. Tao, T. Tan, R.W. Picard (Eds.), *ASCI 2005*, in: LNCS, vol. 3784, Springer, Berlin, Heidelberg, 2005, pp. 350–357.

- [27] V. Francisco, P. Gervás, F. Peinado, Ontological reasoning to configure emotional voice synthesis, in: M. Marchiori, J.Z. Pan, C. de Sainte Marie (Eds.), Proceedings of the First International Conference of Web Reasoning and Rule Systems, in: LNCS, vol. 4524, Springer, Berlin, Heidelberg, 2007, pp. 88–102.
- [28] E. Oren, M. Volkel, J.G. Breslin, S. Decker, Semantic wikis for personal knowledge management, in: S. Bresan, J. Jung, R. Wagner (Eds.), Dexa 2006, in: LNCS, vol. 4080, 2006, pp. 509–518.
- [29] E. Oren, R. Delbru, K. Möller, M. Völkel, S. Handschuh, Annotation and navigation in semantic wikis, in: Völkel Max (Ed.), Proceedings of the First Workshop on Semantic Wikis – From Wiki To Semantics, June 2006.
- [30] M. Buffa, F. Gandon, G. Ereteo, P. Sander, C. Faron, SweetWiki: A semantic wiki, web semantics: Science, Services and Agents on the World Wide Web 6 (1) (2008) 84–97.
- [31] <http://www.w3c.org/RDF>.
- [32] <http://www.oracle.com/lang/es/appserver/index.html>.
- [33] <http://www.w3.org/XML>.
- [34] <http://wordpress.org/>.
- [35] <http://drupal.org/>.
- [36] A. García-Crespo, J.M. Gómez-Berbís, R. Colomo-Palacios, F. García-Sánchez, Digital libraries and web 3.0. The CallimachusDL approach. Computers in human behaviour (2009) (in press).
- [37] S. Bechhofer, F.V. Harmelen, J.A. Hendler, I. Horrocks, D.L. McGuinness, P.F. Patel-Schneider, L.A. Stein, OWL web ontology language reference, Retrieved June 27, 2009, from Sourceforge.net, 2004. Web Site: <http://www.w3.org/TR/2003/PR-owl-ref-20031215>.
- [38] <http://www.w3.org/TR/rdf-sparql-query/>.
- [39] <http://www.openrdf.org/>.



Ángel García-Crespo is the Head of the SofLab Group at the Computer Science Department in the Universidad Carlos III de Madrid and the Head of the Institute for promotion of Innovation Pedro Juan de Lastanosa. He holds a Ph.D. in Industrial Engineering from the Universidad Politécnica de Madrid (award from the Instituto J.A. Artigas for the best thesis) and received an Executive MBA from the Instituto de Empresa. Professor García-Crespo has led and actively contributed to large European Projects of the FP V and VI, and also in many business cooperations. He is the author of more than a hundred publications in conferences, journals and books, both Spanish and international.



Ricardo Colomo-Palacios has been a Faculty Member of the Computer Science Department at Universidad Carlos III de Madrid since 2002. His research interests include Software Process Improvement, Software Project Management and Information Systems. He received his Ph.D. in Computer Science from the Universidad Politécnica de Madrid (2005). He also holds an MBA from the Instituto de Empresa (2002). He has been working as software engineer, project manager and software engineering consultant in several companies including Spanish IT leader IN-DRA. He is member of several Editorial Board Committees of international journals and Editor-in-Chief of the International Journal of Human Capital and Information Technology Professionals.



Juan Miguel Gómez-Berbís is a Visiting Professor at the Computer Science Department of the Universidad Carlos III de Madrid. He holds a Ph.D. in Computer Science from the Digital Enterprise Research Institute (DERI) at the National University of Ireland, Galway and received his M.Sc. in Telecommunications Engineering from the Universidad Politécnica de Madrid (UPM). He was involved in several EU FP V and VI research projects and was a member of the Semantic Web Services Initiative (SWSI). His research interests include semantic web, semantic web services, business process modeling, b2b integration and, recently, bioinformatics.



Francisco García-Sánchez holds a Ph.D. in Computer Science from the University of Murcia. He is a lecturer in the Faculty of Computer Science at the University of Murcia, Spain. He is part of the Knowledge Modeling, Processing and Management Technologies group. He is working on the integration of Agent Technologies and Semantic Web Services. He has taken part in various research projects and has published several papers in JCR journals. He has conducted numerous research stays in some of the most prestigious, Semantic Web and Intelligent Agents concerned research institutes around the world such as the Digital Enterprise Research Institute (DERI), the Centre For Information Technology Research (CITR) and the Stanford Research Institute (SRI).