Valuing Learning Objects shared in an Online Community

ABSTRACT
In this chapter we analyze and discuss how the activity inside a social network impacts on the value of a Learning Object (LO) used in a collaborative e-learning platform. Recent works propose metrics for measuring LO reusability based on a variety of approaches. In this work we combine and extend these approaches in order to design a valuation strategy which helps to identify the usage of LOs inside a social network. Our proposal is to identify the factors that are relevant for the valuation of a LO and determine which of them can be computed automatically from its context of usage, the level of success of its authors and its metadata. Our analysis was performed on a particular social network called LOP (LO Poll) system, which strongly motivates the creation and collaborative valuation of LOs. We present preliminary conclusions obtained from an experiment performed in order to analyze the feasibility of the proposal.

KEY WORDS
E-Learning, Communities, Metadata, Reusability, Quality

INTRODUCTION
For some years now, all of us have started giving more and more importance to the community-society in the information systems. But today, not only does the community consume a technological product, but it has also become part of the product-system. Systems nowadays model the community and its behavior as they learn from the society. Thus, the main idea of enterprise 2.0 is to consider the social software as the element that allows for the creation of new knowledge without being limited to administrating the existing one.

The technology-community link is more important and deep when using technology to support education, since technology plays here a dual role of learning and teaching. On one hand, the e-learning systems must know-infer-learn the behavior and preferences of users (e.g. teachers, students); and on the other hand, they should be the support of the e-learning process. The system learns from the community so as to teach them in the best way.

As we have seen, the value of the knowledge produced in social networks is recognized as a key issue in any organization (Dalkir, 2011; Barão & Rodrigues da Silva, 2011). Moreover, this is extremely relevant in the context of e-learning communities where collaboration and quality of resources is of ultimate importance.

The wide variety of resources used in learning contexts are called Learning Objects (LO) Widley (2001), i.e. pieces of pedagogical information that can be reused. Usually these LOs are stored in repositories of LO, such as MERLOT, ARIADNE, eLERA, Connexions, Agrega, which coincide on the need to create valuable reusable LOs.

In reference to giving value to a LO, we ask ourselves “What is the value of something?” The axiological definition of value according to Frondizi (1992) is related to the value of a socio-personal interpretation according to reality and needs. Therefore, opinions about the LO coming from the community of authors, the community of users and the community of reviewers must be taken into account.

Ochoa (2011) presents a comparative study of LO repositories that show that LO reusability is bigger in repositories in which users are part of a community. Besides this, there exists no standardization about how the activity inside a social network impacts the value of the LOs.
This proposal aims at identifying the factors which are relevant for the valuation of a LO and determine which of these factors can be computed automatically from the LO’s context of usage, the level of success of its authors and its metadata. We perform our analysis on a particular social network called LOP (Learning Objects Pull) system (Dinis & Rodrigues da Silva, 2007, 2009), which motivates the creation and collaborative valuation of LOs. We consider the necessity to prove that the community is important in a measurable way, in the same way that Enterprise 2.0 does in reference to the business performance. Therefore, our proposal measures the value that the LOP’s community gives to the learning objects of the LOP.

We present preliminary conclusions obtained from an experiment performed to analyze the feasibility of the proposal.

RELATED WORK

This section presents an overview of several works which focuses on diverse issues related to the problem of measuring the LO’s quality in terms of reusability.

Factors related to the reusability

Sanz, Dodero & Sanchez-Alonso (2009) proposed a model to evaluate the reusability of LOs which is based on the main following factors:

- **Cohesion** is directly proportional to the semantic density, and is inversely proportional to the amount of: relations, aggregations, concepts and objectives covered.
- **Coupling** is defined by the information given by the relation category of LOM. More relations mean more coupling.
- **Size and Complexity** represent a degree of granularity of a LO (e.g. resource, lesson, course) and are directly proportional to the size, duration and typical learning time of the object.
- **Portability** has a technical and educational aspect. The first one is based on the format of the LO, the requirements of the hardware and software, and the criteria defined in the proposal to assign each of these a value. The educational aspect of a LO’s portability is based on the following metadata: context, typical age range, language, and classification.

These factors agree with fields’ LOM metadata, and are considered according to their importance. The proposal was tested using a set of eLera’s objects, and compared with LORI’s evaluations.

In summary, this work is based on: (1) defining the factors that determined the LO’s reusability; (2) establishing the mapping between these factors and the metadata LOM; (3) specifying the guidelines to qualify each factor.

Further work of Sanz (2010) proposes an automatic process to measure the reusability of a LO from the metrics of the following factors: cohesion, portability and size. The work defines a unique metric which is a combination of the previous metrics weighted according to their importance. In addition, the proposal considers a possible interdependence between the criteria. The validation of the proposal was made through the comparison of the results performed with the Merlot’s and eLera’s evaluations.

In addition, Sanz proposes improving the previous metric with the opinions of users, experts and empirical data such as amount of access, amount of users, and the like.

Work conducted by Meyer, Hannappel, Rensing, & Steinmetz (2007) considers the pedagogical granularity as the most important factor in reference to the reusability. This kind of granularity has been called information’s granularity and it is determined by the didactic functions that an LO may have. A LO may have several pieces of information which have varied didactic functions types (e.g. overview, theorem, example, test). These functions are obtained from ontologies Meder (1999). The research’s goal was to classify the objects’ information in different didactic functions which are defined in the Meder’s ontologies Meder (1999). The function types are: receptive knowledge types (facts,
overview, etc.) and interactive assessments (multiple choice tests, assignment tests). Particularly, they wanted to test multimedia features in the classification process. Therefore, they considered the following features in the classification task: Length of the text according to the amount of words, number of javascript functions, keywords in the page headlines, HTML code containing lists, forms, input elements, choice elements, interaction elements and embedded flash animations.

They constructed a set of objects' information from different sources and multiple authors. Then, these objects were manually classified by reviewers. Four algorithms were used to classify these objects (Bayes, SVM, JRip, C4.5), and the results were compared with the human-made classifications. The experimentation was made several times considering: different levels of details, multi-label classifications, all features and then selected features. Finally, they introduced the possibility of analyzing the position of the objects' information as a relevant factor to consider, besides analyzing the style of speech as a way to identify some didactic functions.

In summary, this work proposes to annotate the objects' information according to didactic functions and to consider these functions as a measure of the LO’s granularity which is related to the LO’s reusability. The correctness of the proposal is based on measuring how close the result is as compared to human evaluation, a fact which is stated in most related works.

Many metrics are based in the LO’s metadata, but are these metadata good enough? In this sense, Ochoa & Duval (2006) proposed measuring the quality of the metadata through metrics calculated automatically. Their work is based in LOM metadata and in the following parameters:

**Completeness** is measured considering that each field of metadata has an importance and a value which will be 1 if the field has information; otherwise, it will be 0.

**Accuracy** is based in semantic distances and content analysis. This comparison between the metadata’s information and the document’s information can be made sometimes by algorithms and others by persons.

**Provenance** depends on the reputation of: the metadata register, the metadata producer, the community, or those who use the metadata.

**Conformance of expectations** evaluates whether the amount of information is enough to describe the LO. The metrics proposed measure the amount of unique information given by each field of metadata. Although there are other aspects related with the conformance, they are covered by others metrics, for instance the completeness of the information.

**Logical consistency and coherence** are based on the consistency rules suggested by LOM about its fields. For example, LOM establishes that if the field structure has the value 'atomic', then the field aggregation level should be 1.

**Timeless** represents the age and the frequency of use of the metadata.

**Accessibility** represents the difficulty to find and understand the information.

In order to carry out the experimentation they had to generate LOM from ARIADNE’s repository (mostly manually) and from OpenCourseWare documents (automatically).

**Environment elements and community behavior**

Some works propose metrics which are based on: the mapping between characteristics and preferences of the environment elements (e.g., situations, users, courses and so on), and the analysis of the users’ behavior patterns. The knowledge about the preferences of the community in the past helps to predict the users’ needs or preferences in the present.

Ochoa & Duval (2008) proposed several metrics used to rank objects in search problems. These metrics can be automatically calculated and they are based on: similarity measurements, frequency of use, and the references between elements (courses, objects and users). These measurements required a historical register of: the users’ behavior, object’s metadata (LOM), and information about the context. The proposed metrics are related to the quality of relevant dimensions Duval (2005):
Basic Topical has to do with the number of times that the objects were selected in queries which are similar to the current query. The similarity is calculated comparing the terms of the queries. Course-Similarity Topical is based on the fact that the objects are used in courses. The proposal considers the level of similarity between the existing courses and the course in which the selected object will be used. Then, these degrees of similarity will be used to sort the output objects according to the courses to which the objects belong. Internal Topical considers the object’s references that are the amount of courses in which the object was used. Rank web page’s algorithms are applied on LO, considering the LO references instead of page links. Basic Personal Ranking uses LO metadata which correspond to previous user selections as the users’ preferences. User-Similarity Personal these metrics are similar to Course-Similar Topical Relevance’s metric, but the similarity is applied to the users. The similarity between two users is given by the amount of objects that they have in common. Basic Situational compares the descriptions of the activities that the learner should be able to do with the LO’s metadata (fields such as descriptions, keywords, and so on). This comparison is made to find the objects which are closer to the objectives. The similarities were calculated according to the terms’ importance and to the semantic distances. Context Similarity Situational proposes that LOs are ranked according to the level of similarity between them and the courses LOs where the evaluating LO are to be used. The similarity is determined by a vector of the frequency of the values of the LO metadata. In summary, this work relies on the need of a historical record of the users’ behavior, and on the analysis of this record to find similarities (between objects, courses, users and queries). The metadata LOM was used. They propose to measure the similarity in different ways: (1) number of relations between elements; (2) semantic comparison between terms and/or values of metadata; (3) number of repetitions of the values of the fields.

Another research, by Yen, Shih, Chao, & Jin (2010), defined a metric to rank LOs. Each object has a value determined by three factors: the time windows in which the object was used, the feedback given by the user (the object was successful or not for the user), and the object download called references (self references and other references have different weight). The importance of the object is made up by its value and the weighted average of the similar objects. This similarity is calculated based on the following LOM metadata: title, language, keyword and coverage of the general category, learning resource type, Intended End User Role, Typical Age Range, Difficulty, and Typical Learning Time of the educational category. The level of similarity is determined by the amount of fields with equal value.

Users valuing and LO’s characteristics-quality
The research by Cechinel, Sánchez-Alonso, Sicilia, and Velazquez Amador (2011) proposed three models to automatic valuing of MERLOT’s resources. These models are based in the LO’s subject (the disciplines of Science, Technology, Mathematics, and Statistics) and materials’ type in a pedagogical sense (Simulation and Tutorial). All material are the same type in a technological sense, they are web pages. The features measured (metrics) were intrinsic characteristics of the web page material, such as: link measures (number of different type of links), text measures (number of words), graphic measures (images, scripts, applets) and site architecture measures (pages: size, number). The LO were classified in good or not-good according to the following algorithms: J48, Simple Cart, PART, Multilayer Perceptron, and Bayesian Network. Then, classification results were compared to the evaluation given by peer-reviewers as a way to test the metrics. This work identifies some features as appropriate to classify LOs as good or not-good material. The features are about material which must be in particular disciplines, type of material and type of format.
The research by Blumenstock (2008) proposed the count of words as a measure of the quality for English articles in Wikipedia. The author came up with 1554 articles selected after a cleaner task, which were used for training and for testing. He compared the results with some other techniques of classification and got good results that supported the proposal. This proposal is simple, clear to understand and easy to measure. In addition, even today, the experts-based evaluations are considered the best approach, so the majority of proposals tend to be validated this way. Besides, many proposals required human intervention during some of their process.

The proposal by Cechinel & Sánchez-Alonso (2011) is strongly based in the valuing given by different kind or users. They consider the ranking given by experts in specific topics and the ranking given by the repository’s users (the community) as well. This research compares both rankings to obtain the similarities about the material’s quality. This analysis searches the best way to use both strategies in the determination of the quality. The experts and users-based rankings are made according to the following dimensions: quality of the content, the ease of use, and the potential of effectiveness as a learning tool.

In general terms, the problem of the human intervention is the time (and consequently cost) required to value the LOs. This is the reason why we need to complement these approaches with automatic mechanisms.

Table 1 shows a comparative analysis of the proposals referred in this section.

<table>
<thead>
<tr>
<th>Research’s Type</th>
<th>Research’s Ref</th>
<th>Main Ideas/Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>J. Sanz Rodriguez (2010)</td>
<td>End metric = F (Cohesion, Size and Portability) \ F / Aggregations = Integral Chiquet (x independence) \ Test = comparison with experts evaluations</td>
</tr>
<tr>
<td>Environment Elements and Community Behavior</td>
<td>X. Ochoa, E. Duval (2008)</td>
<td>History of the users’ behavior, Used metadata LOM, Similarity concepts (objects, courses, users and queries), # ways</td>
</tr>
<tr>
<td></td>
<td>N. Y. Yen, T. K. Shih, L. R. Chao, and Q. Jin (2010).</td>
<td>History register of the user’s behavior, Used metadata LOM, Similarity concept (comparison of the object’s metadata)</td>
</tr>
<tr>
<td>User’s valuing and LO’s characteristics-Quality</td>
<td>C. Cechinel, S. Sánchez-Alonso, M. Sicilia, C. Velazquez Amador (2011)</td>
<td>Identify features classify the LO as good or not-good material, Material = disciplines (math, …), type (tutorial, …), format (web pages)</td>
</tr>
<tr>
<td></td>
<td>Joshua E. Blumenstock (2008).</td>
<td>Quality = Feature = count of word, Material = English article, format (Wikipedia)</td>
</tr>
<tr>
<td></td>
<td>C. Cechinel, S. Sánchez-Alonso (2011).</td>
<td>The rankings of experts and users, Dimensions to rank: quality of the content, the ease of use, and the potential of effectiveness as a learning tool, Compares both to obtain the similarities about the material’s quality</td>
</tr>
</tbody>
</table>

Table 1: Comparative analysis of related work.

ASPECTS OF GIVING AN INITIAL VALUE TO A LEARNING OBJECT

The issue of giving a value to a LO implies establishing which factors are relevant according with a criteria of application and determining how they can be measured. Although the relevant factors might vary according to different realities, there is a consensus in considering the reusability of LOs as one of
the most important ones. Other factors worth considering are: portability, granularity, cohesion, coupling, complexity, adaptability, among others. Many of these factors are interrelated, and all of them affect the reusability. However, in order to determine which factors are relevant, we need to consider the purposes for which we want to give a value to the LO. The value of a LO can be used to solve a variety of problems according to the existing realities: (1) organize the objects in order of importance (ranking problem); (2) search the most adequate object according to user needs (search problem); (3) integrate the LO with others (integrate problem); (4) relate it to a suitable price or cost (business problem); and so on.

Another aspect to be considered to determine the relevant factors is the environment in which the LO is used. For example, the LO may be in connection with other elements such as: social communities (e.g. students, teachers, reviewers), courses in which the LO was used or could be used, and the topics that the LO covers. These connections might be more or less relevant according to the specific reality and the expected use of the LO. For example, in the online community scenario, the reputation and activity of their members is a very important factor to be considered (Iriberry & Leroy, 2009).

The impact of the relevant factors can be calculated in function of the LO’s metadata. We understand by metadata all the information that is used to describe the object, not only social features but also structural ones. Standard metadata seeks to determine which information should be known about the object. For instance, LOM establishes a series of information groups that correspond to different kinds of LO’s information: general, life cycle, meta metadata, technical, educational, rights, relation, annotation, and classification. Another metadata standard is Dublin Core, defined originally to be used in libraries, but currently used in other domains, such as describing web resources. Dublin Core defines the minimum set of metadata necessary to describe resources such as: title, author, copyright. In general, applications use some of the metadata standard but extend or specialize them with specific metadata profiles. The use of standards is important because it adds value in terms of interoperability and reusability. In addition, it may reflect some characteristics of the LO that cannot be drawn from the analysis of the object itself.

Another important aspect to consider in the problem of valuing a LO, is valuing each of the factors identified as relevant from a technological perspective but also from a socio-pedagogical one. These factors are determined by LO’s metadata; therefore, these metadata must be also composed by technological and pedagogical features. Moreover, another perspective to be considered for the LO analysis is the LO’s structure, content and environment. For example, Figure 1 shows a diagram of the analysis of the complexity factor, which is analyzed by examining the structural, content and environment features through a technological and socio-pedagogical point of view. Following Sanz (2010), the features are grouped in two categories: Structural and Content, but considering our analysis of relevance we also added the Environment grouping category. The size and format are some features from the Structural group whereas words and pages belong to the Content. Some features may be considered in both perspectives simultaneously. As shown in Figure 1, the duration factor, i.e. the time that the user needs in order to hear or view the material, is considered as a technological feature (depending on the hardware device that supports the user’s experience) but it is also considered as a pedagogical feature (depending on the level of concentration required by the topic). Technological and Pedagogical classifications are presented in the works of Sanz, Dodero. & Sanchez-Alonso (2009), Sanz (2010) among others.

When analyzing the complexity factor in relation to the Environment, the following features are identified: the level of the audience for which the LO is designed (e.g. courses, students), the depth of the topics it deals with (e.g. basic, medium or advanced), and the type of courses in which the object can be used (e.g. online, classroom, or semi-presenciales courses). Some of these features can be also analyzed either from the technological or pedagogical point of view.
In summary, to measure the value of LOs in an application we need to establish: what the relevant factors are (such as complexity, portability or granularity), what the point of view from which we would like to analyzed them is (structural content, environment, technological, educational, or any combination of these); which the needed metadata to describe these factors are and whether they follow some standard (e.g., Dublin Core, LOM); how the metadata values are loaded (automatically or manually); and, finally, how the factors are measured.

THE LOP SYSTEM

The Learning Objects Pool (LOP) System is a collaborative Web-based application that supports user communities as part of the teaching and learning process (Dinis & Rodrigues da Silva, 2007, 2009). Technically the LOP system is a LO repository that promotes the active collaboration of its users, both in submission of LOs, as well as in the submission of relevant information that will contribute to its understanding. The system distinguishes and rewards users who contribute the most in the creation of quality LOs. In addition, LOP also compensates the end-users active and regular collaboration through a credit-based system that calculates the level of users’ collaboration. This mechanism also updates the LO value according to its popularity, similar to the stock exchange metaphor, where the LO value raises or falls depending on the number of acquisitions which represent the LO’s popularity.

The LOP system currently provides a good stability in terms of usability and functionality, and has been instantiated for different usage scenarios, namely the BOA-GPI and the VemAprender.net instances. The BOA-GPI” is supporting the “IT Project Management” course at IST (Technical University of Lisbon), supporting several generations of students and faculties assigned to this course. The VemAprender.net” is an open and public-access instance, for everyone interested in sharing or reusing LOs, particularly targeted to students and teachers at K-12 education levels. (There is also the “BOA4OpenCommunity”\textsuperscript{viii}
just for simple demonstration and evaluation purposes.) In the context of this research we have used the data available from the VemAprender.net instance.

Teachers, students and administrators work collaboratively in this LOP community (the VemAprender.net platform) sharing and reusing LOs to support the teaching-learning process. This LOP repository not only has the LOs, but also an important set of information – LO’s metadata – that describe them, such as topics, type of resource, description, language; in addition to registers of the users and information on activities (e.g. submitter, comments, suggestions).

The LOP system aims at achieving LOs excellence through an open and competitive environment. This platform allows the centralized gathering of LOs besides their reuse in different courses. LOs have an associated value and the users can download them if they have enough credits. Users increase their own credits by doing different activities such as creating and sharing LOs or commenting, rating, describing instructions or good practices related the LOs or even by writing suggestions for new LOs.

Currently, when a user submits a LO to the repository he first has to upload the LO and its metadata and he has to specify an initial value to that object. Then, optionally, a reviewer can evaluate the object being able to comment and give recommendations to the author and/or negotiate with him the initial value of the object. Finally, the LO becomes available to the community members who might download it. The value of the object varies periodically, increasing or decreasing, according to the number of downloads carried out. Figure 2 shows the LO submission workflow.

![LO submission workflow](image)

*Figure 2: LO submission workflow (Dinis and Rodrigues da Silva, 2009).*

The LO can have more than one author. Part of the information that the submitter should introduce is related to the LO’s authors, their percentage of involvement, and which the responsible author should be (the LO responsible is the only one who is enabled to change some information about the object such as its metadata).

Additionally, the LOP system should keep track of other information to better categorize and organize LOs such as groups, level of audience and size of the object among others. This information is encoded in different types of metadata, called intrinsic or manual metadata. The metadata that LOP system employs to describe its objects are made up by 39 fields, some required and most of them optional, which extend the Dublin Core Standard.

The LOP allows for LOs to be described by the following metadata groups:

**General Information** group is the main information which is required in most fields. In this group of metadata we can find the following information:
• **Submitter:** Corresponds to the information about the author who submits the LO. Each LO has to have at least one author. If the LO has more than one author, authors and level of participation in the development of the LO should be assigned in the *Authoring metadata* group.

• **Group and Level of audience:** This information describes the object’s profile from a pedagogical perspective. On one hand, the Group corresponding to the context where the LO may be used. This information is specified by the submitter. Each user of the LOP system is associated to groups for which he is able to publish, consult and download LOs. Moreover, if the object can be used in more than one group, the detail can be stated in the *Topic*. On the other hand, *Level of audience* indicates whether the LO is designed for students in basic, intermediate or advanced levels.

• **Title, Description, Keywords, Language, Date of creation and Format:** This information depict the features of the LO itself. Though some of these metadata could be uploaded automatically analyzing the object’s content file, the information is currently uploaded manually by the object’s submitter.

• **Initial and Minimum value of the object:** This metadata indicates the initial and minimum value of the LO according to the criteria of his author.

**Topic group** permits to indicate the topics that best describe the LO; each group has its specific hierarchy of topics.

**Authoring group** is the information about the users that participate in the authoring of the object.

**Collaborators group** is information regarding people who in spite of not being users of the system or authors of the objects, have collaborated somehow in its design or production.

**Relationship group** is the information that states relationships between the LO and others LOs in the repository; several types of relationship are supported, such as: part-of, version-of, addition, or generalization-of.

**Types group** is about the type of material from a pedagogical perspective, for instance: exercise, simulation, example, and others.

**Images group** includes images associated to the LO, those images that appear in the content of the objet itself as well as other suggestive images.

**Other Information**, this group contains other information such as author’s rights, editorials, instructional methods, coverage or the source of the object.

The proposal discussed in this paper aims at using the information specified by the submitter, as well as information extracted automatically from the object, to suggest the initial value for the corresponding LO.

---

**THE INITIAL VALUE OF A LEARNING OBJECT IN THE LOP SYSTEM**

Based on the related work and the characteristics of the LOP systems we identify the following main factors which determine the LO’s capacity of reusability: granularity, complexity, portability, completeness and provenance of metadata.

**Granularity:** We understand the granularity as the number of topics covered by the LO. We consider the criteria used in (Meyer, Hannapel, Rensing & Steinmetz, 2007): “Producing large units (e.g. whole courses) provides the best usability in the first place, but decreases the probability that this large unit be re-used in another context”.

**Complexity:** We consider the complexity of a LO as the difficulty to understand the LO. In this sense, we believe that the more information the LO has, then the more understandable the LO is. In this work, we associate the measure of the complexity to the measure of the amount of information according to pages/slice, words, paragraphs, images and formulas that the LO contains. However, the relation between complexity and amount of information depends on the type of file of the LO and the context of its use.

**Portability:** in terms of reusability the capacity of carrying an object from one context to another (courses, platforms, and users) is very important and it is necessary to analyze the pedagogical and technological features of the object to measure these capacities. From the technological point of view the format of the
object must be considered. Some optional metadata should be taken into account but we don’t use optional features in the proposal. From the pedagogical point of view the granularity and the complexity determine the portability. And we consider the context for which the author creates the object is given by the group features of LOP.

**Completeness**: is based on the amount of optional metadata of the LO filled by the user. The relevance that we assign to this factor is in order to encourage authors to load the optional metadata.

**Provenance**: like Duval (2006) and Stack Overflow community we consider the source’s reputation as a relevant factor. In the LOP system the optional metadata is provided by the authors; consequently, we measure the authors’ reputation.

These factors are measured through the properties of LOs (structural and content features), users and context. Therefore, our proposal consists of determining the value of a LO in terms of the following elements: Authors, Context, the content of the LO, the LO’s Metadata and the LO itself (Figure 3).

![Figure 3: Assessment Process of LOP Objects.](image)

The proposal follows the simplicity principle. The aim of the functions proposed to calculate the LO’s value is to be simple to understand, implement and process.

We determine the initial value of a LO taking into account: the value of the author, the value of the context, and the value given by the intrinsic characteristics of the object. We propose that the value of the LO can be calculated as the sum of the values of the elements mentioned above (in order of importance) as shown in the following formula:

\[
V_{LO} = \alpha_A V_{AUTHORS} + \beta_C V_{CONTEXT} + \mu_{LO} V_{LO_{INSTRINC}}
\]

\[t.q. \ 0 \leq \alpha_A, \beta_C, \mu_{LO} \leq 1 \quad \alpha_A + \beta_C + \mu_{LO} = 1\]
Where:

$V_{AUTHORS}$ represents the reputation value of the LO’s authors which is associated not only with their profile but also in relationship with their level of participation in authoring tasks and with the level of success obtained.

$V_{CONTEXT}$ represents the importance of the context of usage and the pedagogical granularity of the LO. The first is defined by groups and subjects in which the LO can be used, and the second is the level of LO’s specialization.

$V_{LO_INTRINSIC}$ represents the values of the intrinsic properties of the LO. Some of those metadata can be automatically extracted which determines the LO’s portability, and other metadata are optional and are related with the completeness factor.

The relevance assigned to the $V_{AUTHORS}$, $V_{CONTEXT}$ or $V_{LO_INTRINSIC}$ in the calculation of the initial value of the LO is specified by the constants $\alpha_A$, $\beta_C$, $\mu_{LO}$. In our scenario of LOP system, we propose to give the most importance to the intrinsic characteristics of the object because of the mechanisms of loading used, then the characteristics of the authors, and last the context. In the following subsections we present our proposal of formulas to calculate the values of Authors, Context and Intrinsic Metadata; $V_{AUTHORS}$, $V_{CONTEXT}$ and $V_{LO_INTRINSIC}$ respectively.

### The Authors

Similarly to the way StackOverflow uses the concept of reputation, the LOP System encourages the users to create reusable LOs. Therefore, we consider measuring the reputation in relation with the users’ participations (authoring task). We propose measuring the value of an author not only by the level of activity of the author, but also considering the level of acceptance of his LOs among users of the community (success), and the profile of the author (e.g., student, teacher,). Therefore we consider the $V_{AUTHOR}$ as a function of the initial value of the author’s type plus the amount of LOs published by the author in the repository ($N_{LO_PUB}$) and the value of success of his/her LOs (distinguished by LOs download at least once and those which present a good success of downloads: $N_{DOWNLOAD} + V_{LO_DWN}$). This is depicted by the following formula:

$$ (2) \quad V_{AUTHOR} = V_{AUTHOR\_TYPE} + N_{LO_PUB} + N_{DOWNLOAD} + V_{LO_DWN} $$

The value of the author $V_{AUTHOR\_TYPE}$ is initialized with a fixed value depending on the author type configured by the system administrator. In the beginning, the importance of the author’s group and the LO’s topics encourage the community to download the object, but only the author’s profile is measured. The LOP system has four type of authors: students, teachers, students’ representatives (e.g., parents), and others. The teacher is the most valuable type.

The level of activity of the authors ($N_{LO_PUB}$) influence directly in the value of the author because of LOP system strategy to motivate the participation of the members of the community to produce LOs.

The success of the author is measured through the number of downloads of his/her LOs. Each download is added to the author’s value ($N_{DOWNLOAD}$) and this value is increased with the average of the LO’s value that have had a special success ($V_{LO_DWN}$), this is, those LOs which have been downloaded at least a minimum number of times, for example the 10% of the number of members of the community.

$$ (2.1) \quad V_{LO_DWN} = \frac{\sum_{i=1}^{N} V_{LO}(O_i)}{N_{LO_PUB}} \quad \text{t.q. } N_{DOWNLOAD}(oi) > k, \quad k = 10\% N\text{-members} $$

Finally, the value of a LO as compared to the value of its authors ($V_{AUTHORS}$) is the average of the authors’ values ($V_{AUTHOR}$) affected by the percentage of participation that each author has in the authorship.

### The Context
In LOP system, when the LO is uploaded to the repository the author can indicate the Groups and the Topics in which she/he considers it is most appropriate to use the LO. Even though the groups have topics, the author should indicate the specific topics covered by the LO (selected topics).

The Group and the Selected Topics determine the profile of the context in which the LO can be used, therefore the proposal considers the group’s value and the selected topics’ value in the valuation of the context.

As mentioned above, we consider that a fine granular LO is more valuable than general ones. This is because the greater the amount of topics covered by an object, the less portable and reusable this object is. Then, we consider that the number of topics covered by the object is a measure of the LO’s reusability and we propose the following formula to calculate the value of the context of a given LO:

\[ V_{\text{CONTEXT}} = \alpha G \times V_{\text{GROUP}} + \beta T \times V_{\text{TOPICS}} \quad \text{where} \quad 0 \leq \alpha G < \beta T \leq 1 \quad \alpha G + \beta T = 1 \]

The initial value of the groups and topics are configured by the system administrator according to the application. But for the topics we propose the formula (3.1) because LOs with only one topic are the most specialized object. Then, they have the highest level of reusability and they are the most valuable. When the LO has different topics, we propose that the value of the topic corresponds to the less valuable topic in relation with the total amount of selected topics for the LO.

\[ V_{\text{TOPICS}} = \begin{cases} V_{\text{SELTOPIC}} & \text{one topic selected} \\ V_{\text{SELTOPIC.MIN}}/N_{\text{Topics}} & >1 \text{ topic selected} \end{cases} \]

Although the proposal allows for a different weight to the group and the topics (\(\alpha G, \beta T\)), we consider that the information given by the topics is more important because it refers to the knowledge transmitted by the LO. Moreover, the value of the group is general information about the context of the LO that depends on the application, i.e. a large number of members in the group may be considered more or less valuable.

**LO Content and Metadata**

The proposal aims at estimating the value of the intrinsic characteristics of the LO \(V_{\text{LO_INTRINSIC}}\) described by metadata. Some metadata can be obtained automatically \(V_{\text{LO_AM}}\), while others are given manually by the author \(V_{\text{LO_MM}}\). Then we propose to calculate the value of the intrinsic characteristics of the LO as follows:

\[ V_{\text{LO_INTRINSIC}} = \alpha MM \times V_{\text{LO_MM}} + \beta MA \times V_{\text{LO_AM}} \]

Where \(0 \leq \alpha MM \leq \beta MA \leq 1\) \(\alpha MM + \beta MA = 1\)

The proposal gives different importance to manual metadata and automatic metadata \((\alpha MM, \beta MA)\) according to how they are obtained. The automatic extraction is a reliable mechanism; therefore, we consider that the information automatically obtained is more valuable than the information given by the submitter \((\beta MA > \alpha MM)\). In the next sections we describe: how to calculate the value of a LO given by manual and automatic metadata, the LO’s features considered; the criteria used to assess the metadata that the user must sign in; and the criteria used to assess the metadata automatically obtained.

**Automatically Extracted Metadata**

Automatically extracted metadata corresponds to characteristics of the LO such as: format, size, number of pages, date of creation, and others. However, deciding which part of these metadata is relevant for each LO depends on the application and on the type of the LO, i.e. if we consider a presentation LO we would take into account the number of slides, rather than its size. On the other hand, if we consider a video LO it
would be of interest the duration of the LO, rather than its number of pages. In this work we focus on text and presentation files.

We consider the format of a LO file as the level of technical portability of the object, in the sense of the actual degree of diffusion (knowledge and use) of the type of file. For example, a PDF file is managed by almost all PCs, while the DOCX file is a commercial format and not all PCs can read this kind of file; consequently, PDF is a more portable format than DOCX. Thus, the proposal in order to value this feature (format feature) is to assign values to the different MIME format. The actual values for each format may be established by the administrator of the system at configuration time.

Another issue about the LO content is the amount of information that the LO possesses. The amount of information can be seen as a measure of how in detail the issues have been covered. In this sense, the more the information the object has, the more valuable it is, because it may be less complex to understand its content. Consequently, it may be easier to use it in a variety of contexts, a fact that makes it more pedagogically portable and reusable. In our context, words are atomic elements that contain some information, and paragraphs may be considered more relevant than words since it is expected that each paragraph transmit an idea or a unit of knowledge. The same applies to images and formulas, since we can assume that they transmit an idea. Therefore, within these considerations, in order to value the properties of a LO that we can obtain automatically we distinguish between the content and the format of the LO and the content between information and ideas.

\[
\begin{align*}
V_{DA-AM} &= V_{Content} + V_{Format} \\
V_{Content} &= V_{Info} / N_{PageORslices} \\
V_{Info} &= V_{OneInfo} \times N_{Words} + V_{OneIdea} \times (N_{Paragraphs} + N_{Images} + N_{Formula})
\end{align*}
\]

Where \(1 \leq V_{OneInfo} < V_{OneIdea}\)

The values for information and idea \((V_{OneInfo}, V_{OneIdea})\) must be configured by the administrator of the system according to the specific application.

**Manual Acquired Metadata**

There are some LO’s properties which are very difficult to be automatically extracted, for example, the pedagogical objective of the LO, the appropriate audience to use the LO, and so on. Most of this information must be given by a person who knows the LO, we call this information Manual Metadata. This section presents the manual metadata used by the LOP system and how they participate in our valuation of a LO.

Optional metadata completeness is an important aspect because they reflect the concern of the author to describe well her/his LO. And giving value to this factor motivates the author to provide this information. Table 2 shows the LOP’s groups of metadata and indicates the names of the factors where they are used in order to value the LO.
Table 2: “Classification of Metadata Groups”

<table>
<thead>
<tr>
<th>Metadata Group</th>
<th>Description</th>
<th>Required/Optional</th>
<th>Group’s Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Information</td>
<td>Keywords, description, title, group, submitter, creation date, format name, language, level of audience, and so on.</td>
<td>Required</td>
<td>--</td>
</tr>
<tr>
<td>Topics</td>
<td>Topics of the LO</td>
<td>Optional</td>
<td>V%M_Topics</td>
</tr>
<tr>
<td>Authoring</td>
<td>Authors and their degree of participations</td>
<td>Optional</td>
<td>V%M_Auth</td>
</tr>
<tr>
<td>Images</td>
<td>Images, give an complementation information</td>
<td>Optional</td>
<td>V%M_Imgs</td>
</tr>
<tr>
<td>Types</td>
<td>LO’s type, e.g. exercise, example, simulations, and so on.</td>
<td>Optional</td>
<td>V%M_Type</td>
</tr>
<tr>
<td>Collaborators</td>
<td>Collaborators that aren’t BOA’s users</td>
<td>Optional</td>
<td>V%M_Colab</td>
</tr>
<tr>
<td>Relations</td>
<td>Relation with others LO, such as: is part of, is version of, is referenced by, and so on.</td>
<td>Optional</td>
<td>V%M_Rel</td>
</tr>
<tr>
<td>Others</td>
<td>Information about coverage, copy rights, method of instruction, publisher, and others.</td>
<td>Optional</td>
<td>V%M_Others</td>
</tr>
</tbody>
</table>

In order to calculate the intrinsic value of the LO, we consider the level of completeness achieved by the author in the description of manual metadata for each group ($V_{M\_TOPICS}$, $V_{M\_AUTH}$, etc.).

(4.2) \( V_{OAIntMM} = V_{M\_Topics} + V_{M\_Auth} + V_{M\_Imgs} + V_{M\_Types} + V_{M\_Colab} + V_{M\_Rel} + V_{M\_Others} \)

The following experiment section details the completeness values suggested for each of group and the criteria used to assign them.

EXPERIMENTATION: VALUATION PROCESS OF LOP OBJECTS

A preliminary test was implemented to the metrics proposed to value the LO’s LOP. This test was made to show the feasibility of the metrics and to identify the conditions in which a completed test must be done. A completed test must be carried out in order to compare the results with human experts valuation, and with the success of the object in LOP over the time. The proposal is based on the value and the weight of some elements and features, the criteria to configure these values are described in this section. In addition, the experimentation and its results are described too.

Configuration of the Criteria’s Values

The table 3 specifies the criteria, conditions and suggested values to assign the weights required by the formulas (1) (3) (4).

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Conditions</th>
<th>Suggested Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>LO’s intrinsic features are more important than the reputation’s authors.</td>
<td>$1 &gt; \mu_{LO} &gt; \sigma_{A} &gt; \beta_{c} &gt; 0$ and $\mu_{LO} &gt; \sigma_{A} = \beta_{c} = 1$</td>
<td>$\mu_{LO} = 0.5$</td>
</tr>
<tr>
<td>Context’s importance is the less relevant</td>
<td>$1 &gt; \beta_{c} &gt; \sigma_{A} &gt; 0$ and $\sigma_{A} = \beta_{c} = 1$</td>
<td>$\sigma_{A} = 0.4$</td>
</tr>
<tr>
<td>Topics are more important than Groups</td>
<td>$0 &lt; a_{STM} &lt; a_{STM} = 1$ and $a_{STM} = a_{STM} = 1$</td>
<td>$a_{STM} = 0.4$</td>
</tr>
</tbody>
</table>

Table 3: Weigh’s Configurations

Because of formulas (2) (3) and (4.2) the value of the following elements and features must be configured, they are: authors’ type ($V_{AUTHOR\_TYPE}$), groups ($V_{GROUP}$), topics ($V_{TOPIC}$), the value given to different file types ($V_{Format}$), and the value of each optional metadata groups ($V_{M\_TOPICS}$, $V_{M\_AUTH}$, $V_{M\_IMGS}$, $V_{M\_TYPES}$, $V_{M\_COLAB}$, $V_{M\_RELS}$, $V_{M\_OTHERS}$).
The value, by default, of an author that does not have LO in LOP yet is according to the type of authors. The value of a teacher set to 4, to a student is 3, to teaching’s responsible is 2, and to others is 1.
The criteria used to assign the groups and topics’ value was: if there are no references about the group or topic then its value must be 1 (default value), if there are some references then the value must be 2, but if there are several references the value suggested is 3. These values will be assigned by the person who loads the new group/topic. The value assigned in each group of optional metadata was: 40 to the groups that are more important in the calculus process which are VM_TOPICS, VM_AUTCH; 30 to VM_TYPES and VM_RELS; and the least important groups with the value of 20 to VM_imgs, VM_COLAB and VM_OTHERS. The end values will be a percentage of these values according to the degree of completeness (normalized). Regarding the elements that can be identified in the content of the file, the proposal is to consider the word as the atomic piece of information (V_OneInfo) and its value will be 2, and consider one image or paragraph or formula as one basic idea to transmit (V_OneIdea), the value of one idea will be 3.
To value the type of file (V_Format) the criterion defined is that the not generalized/diffused format must be valued 1, the diffused format which is not free/open must be valued 2, and the diffused and free format must be valued 3. In our case we tested with DOC, PPT and PDF format because they are the most used in the LOP system, for which we used 2, 2 and 3 values respectively.

**Experimentation**

We performed an experiment within VemAprender.net instance of LOP System (Dinis P., Rodrigues da Silva A., 2007) which had: 7 groups, 53 authors, 163 LOs. This preliminary experiment aimed at detecting problems in the data used, the limitations of the platforms and metadata, and to identify the indicators which needed to be modified to produce a great impact on the evaluation metrics of the LO. In the following section, we analyze each of the considered indicators.

**Authoring**

We simplified the experiment assuming one author by object. The value of reputation of each author was calculated according to the formula (2). Therefore we used the values of the author’s type and the LO; we used the amount of objects by author, the amount of downloads by author, and the amount of downloads by object. The results were analyzed with the purpose of verifying the existence of a relation between the calculated values and the behavior that the community has had. We analyzed the authors of the most successful LOs (and also the least successful ones) with the best (and worthy) authors according with the proposed formulas. None of the worth LOs had been authored by one of the best authors, and the same situation occurred with the best LO respect to the worth authors.
The analysis was made over general terms because the calculations used initial LO values assigned by the author, consequently these values don’t have a unique and objective criterion. Therefore the set of authors and LOs considered were filtered. The authors which don’t have at least three download over their objects were removed. Two authors were removed because they have objects which were initialized with very low or high value compared with the rest of the object initial values. Some authors had low rate because they were in LOP recently. These cases were identified by a human analysis. LOP had 53 authors and finally only 15 authors were used in the experimentation.

We run the experiment with 15 authors selected using the previous criteria, each with their respective values of authorship (range of values between 107 and 263), and value of success (the amount of downloads on their LOs). These authors are classified in three intervals: great success, medium and insufficient success. The process obtained 5 authors in each category. None of the authors of recognized success was rated negatively nor was any author misclassified as insufficient success. Furthermore, the analysis of the cases assessed as medium by the process shows that this result depends on the initial value assigned by the author of the LO. This initial value is observed to be very different among the various members of the community; therefore, a possible way to deal with this deviation is assigned reviewers...
with a uniform view of the community or automatically produce a valuation based on intrinsic LO’s metadata.

**LO Metadata**

As it has been analyzed in previous sections, there are a great amount of LO’s features specified by metadata that are relevant factors in the valuation of an LO. Some values of these metadata can be automatically extracted, while others must be given by the author himself/herself.

Regarding the first group of metadata, we propose a metric that represents their impact in the LO’s value which is depicted by formula (4.1). In the next paragraphs we use this formula to evaluate the LO’s value inside our experiment.

Step1: the LOs were grouped according to the files type PDF (15 LOs), DOC (4 LOs) and PPT (29 LOs). The LO were categorized according to two criteria: the level of success (high, medium, low), and the value of the LO (high, medium, low).

- We analyzed the LO categorized simultaneously as ‘most successful and less valuable’, and as ‘less successful and most valuable’. The objective is to find the reasons of the inconsistency.
- We analyzed the LO categorized as ‘half successful and half valuable’. The objective is to identify the characteristic of the general behaviors

Step2: the most successful LO are analyzed. We compare their level of success with the amount of: images, formulas, pages, paragraph and words that its files have.

Step3: the LOs with similar levels of success but with different types of file are compared.

The early results show the need to have values given by some mechanisms which are reliable and more independent of the environment factors such as context and authors. We could compare the calculation values with values given by experts. In this preliminary test the LO’s success (amount of downloads) was used in some analysis. Although the LO success gives an idea about the value given by the LOP community, this success is strongly influenced by environment factors such as authors, times, and context. For example, we find many objects with low values and high level of success because of the authors’ reputation; while other objects had high values and low level of success because of their availability time in the platform. Other erratic indicators appears when the metric to valuing the LO is based in some metadata that is not possible to be extracted for the specific LO's format.

In order to refine the corpus of the experiment each erratic LO was eliminated and other LOs that fall into the same situation where added to repeat the test and identify common factors in the general case.

The optional metadata gives pedagogical information that can not be obtained if it is not through the author. Therefore, it allows for evaluation of pedagogical aspects of the object which are relevant, but on a first stage our objective is to motivate the author to load this information. This is the reason why on this proposal the completeness of the metadata loaded by the author will be considered as a priority. Measuring the quality of the information given by the author and the value that each of these data gives to the object is out of the reach of this research. Therefore, the metrics proposed (formula 4.2) measure the level of completeness (the amount of fields filled by the author). An important aspect of this kind of information is the identification of the necessity to add metadata to the LOP system.

**Context**

The metrics to measure the value of a LO’s context is given by the formula (3). These values are calculated considering the configured values given to the groups and topics, and the amount of topics covered by the LO. However, in our scenario we find that most of the LOP’s LOs had the group of the author by default, and did not have topics information. Another unexpected problem that appeared in the experiment was the difference between the level of specificity of the LO’s topics. Some LOs have only one topic e.g. “Introduction to TICs”, while others had various topics with different level of generality,
for example ‘informatics’, ‘TICs’, ‘programming’. This kind of difference affects the metrics because the proposal considers the amount of topics selected by the author.

**Preliminary Results**

These preliminary results led us to determine conditions to the next complete experimentation; and let us identify possible extensions to the platform and metadata that would improve the results. They are:

1. The complete test must have a set of LOs with an expert evaluation to compare the calculated values with these evaluations.
2. The preliminary test verified that the authors’ reputations may determine the value of the intrinsic features of the LO. Therefore, one conclusion is that in the complete test we would isolate the author’s reputation to evaluate the other factors. Therefore, we suggest that the LOs should be submitted by one unknown author.
3. We identify some problems related to the age of the LO and the authors. A LO can have few downloads because it is a new LO and not because of its quality. The analog situation occurs with the authors old in LOP. Therefore, we suggest to submit the set of LOs to the experimentation at the same time by a new LOP user/author.
4. Regarding the context value results, we observe that the capture of the topics’ information doesn’t have clear criteria on how to fill it. As our proposal measures how specific the LO is according to the topics, then the calculation value is affected. Therefore the platform must be clear about the criteria to capture this metadata. One possible solution may be to use taxonomy of the topics group presented as a tree to the user, who must select the topics.
5. The test carried out verifies that one of the most important elements in the re-use of the object is the community. Therefore, it is an added value to know and consider who re-used the LO. Consequently, the registration of the affiliations of the users that make downloads could improve the estimation of the LO’s value and the authors’ value.
6. By adding the pedagogical objective of the LO, the measure of granularity and portability from the educational point of view could be improved. For example, one possibility is to describe them through verbs associated with Bloom’s Taxonomy (Anderson, L.W., and D. Krathwohl, 2001).
7. Finally, we have to consider some limitations that the extractor of metadata tools may have. Tools used to extract the metadata used in the valuing process could affect the conclusion to which we arrived.

**FUTURE RESEARCH DIRECTIONS**

Modern systems learn from the community in order to give assistance to the community in the best way. On one hand the e-learning system must know-infer-learn the behavior and preferences of users (teachers, students, etc.), and on the other hand, they should be the support of the e-learning process. In this sense, we find different types of activities which are developed today: connection repositories, hybrid collaborative recommender systems based in the user’s opinion and in the analysis of the community behavior, and the possibility to integrate the produced data in order to produce knowledge as it is performed in the linked data approach.

**CONCLUSION**

Research on measuring the LO’s quality according to the reuse has been presented and factors which are related to the reusability and their relation with LOM have been described. We have stressed the needs of the metadata. Research which measures the reusability based on the type and topics of the material has been explored. The importance of the community in the e-learning systems and in the metrics to measure the success of a LO was presented, not only as mentioned by recent research but also in our proposal. The problem of valuing a LO was analyzed from different points of view: technological, socio-pedagogical, structural and content. The LOP systems were described and the problem of giving an initial value to their LOs was analyzed, moreover, a solution was proposed and tested.
The main contributions of this work are the following:

- It is confirmed that LOs are as important as the community itself. This means that the user is important both as an individual and as part of a group. In our case, the experiments show that the object’s success is strongly determined by the author’s reputation.
- We define and discuss an automatic process to calculate the initial value of LOs, based on explicit but also implicit LO’s features.
- We also propose an automatic process to measure the implicit/intrinsic features of a LO.
- We show that although the calculation process of the LO’s value can be automatic, the experts’ opinion and the community’s opinions affect directly the valuing of those LOs. These opinions and feedback should be also considered at the initial stage but also periodically to allow updating the value of such as LOs and authors.
- We also propose a framework of metrics that has into consideration motivational strategies to promote the users’ participation in the community.

ACKNOWLEDGMENTS
This project is carried out with the support of the SOLITE-CYTED and LACCIR_R1210LAC007.

REFERENCES
Augsburg. Germany: ACM


**ADDITIONAL READING**


**KEY TERMS & DEFINITIONS**
DC: metadata standard to describe resources
E-Learning: systems which give support to the teach task
LO: resource, material used to teach.
LOM: metadata standard to describe learning objects
Metric: measure of something
Ranking: relevance value
Reusability: the capacity of adapting to be used in different ways
Value: quantification of the importance

INDEX
Community, 11
context, 6
Context, 11
E-learning, 17
Factors, 2
Granularity, 9
Learning Object, 1
Metadata, 3

Metric, 3
Portability, 9
Quality, 2
Reputation, 10
Reusability, 2
Value, 1

\textsuperscript{i} http://www.merlot.org
\textsuperscript{ii} http://www.ariadne-es.org/
\textsuperscript{iii} http://www.elera.net
\textsuperscript{iv} http://www.cnx.org
\textsuperscript{v} http://proyectoagrega.es
\textsuperscript{vi} http://isg.inesc-id.pt/BOA-GPI
\textsuperscript{vii} http://vemaprender.net
\textsuperscript{viii} http://isg.inesc-id.pt/BOA
\textsuperscript{ix} http://www.dublincore.org/
\textsuperscript{x} http://stackoverflow.com/