

Open Access to Mexican Academic Production

Silvia I. Adame*, Luis Llorens

Institute of Engineering, University Autonomus of Baja California, Mexico

Copyright © 2016 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Abstract This paper presents a description of the metadata harvester software development. This system provides access to reliable and quality educational resources, shared by Mexican Universities through their repositories, to anyone with Internet Access. We present the conceptual and contextual framework, followed by the technical basis, the results and future work. This paper is based on the experience gained from working with the technical committee of the project sponsored by CUDI-CONACYT titled: Metasearch of Educational repositories to Promote the Use of Learning Objects and Open Educational Resources: Best Practices.

Keywords Mexico, Metadata Harvester, Digital Repository, Open Educational Resources, Search Engine

1. Introduction

The search for information on the Web is an everyday activity, finding free, reliable and quality information is a challenge. Most of our search results are widely varying quality, and it is difficult to find reliable educational content online.

According to estimates by Van de Sompel [1], the Web is growing at an amazing rate; every minute over 70 new domains are being registered and more than 500,000 documents are being added to websites. In Mexico, today, there are more than 4.1 million people that access the Internet on a regular basis [2]

Rapid Web expansion and increasing Internet access bring great opportunities and challenges for Universities; among them, the opportunity to develop a culture of sharing and reusing scientific, academic and cultural information to benefit those with internet access. And the challenge to disseminate those digital content to be reached by Web users.

Adame *et al* [3] explain, that recently in Mexico, the Open Access Movement which promotes to use information technology to help equalize the distribution of knowledge; has triggered the development of OER, and the implementation of repository systems in universities.

This paper research, aims to show the development of a

meta-connector that allows harvesting various repositories that can be used by infomediaries, in order to facilitate the task of finding, evaluating and sharing high quality content to all Internet users who know it.

This is based on the experience gained from working with the technical committee of an interagency project sponsored by CUDI -CONACYT, titled: “Metasearch of Educational repositories to Promote the Use of Learning Objects and Open Educational Resources: Best Practices, and the research “Characterization of Mexican Educational Repositories”

2. Conceptual Framework

In 2002, the term Open Educational Resource (OER) was coined by the United Nations Educational, Scientific and Cultural Organization (UNESCO), to refer to educational resources generated to provide digital access through information and communication technologies (ICT), to be used for non-lucrative purposes, following the guidelines of Open Access [4]

The term OER is largely synonymous with the term OpenCourseWare (OCW), although the OCW is defined as a free and open digital publication of high quality university-level educational materials.

The William and Flora Hewlett Foundation defines *OER as* “resources for teaching, learning and research resources that reside in the public domain or have been released under a licensing scheme that protects intellectual property and allows its use as free and the generation of derivative works for others” [5]

OER are identified as, course materials, modules, books, videos, tests, software, and other tools, materials or techniques used to support knowledge access.

The OER have a transformative power that lies in the ease with which such digitalized resources can be shared via Internet. One key differentiator between OER and any other educational resource is license. An OER incorporates a license that facilitates the use, reuse and potential adaptation, without first requesting permission from the copyright holder [6]

One vital aspect of the OER economy is the role of

metadata. Metadata is often simply defined as “data about data”. The OER need metadata or tags that allow them accessibility, reusability and interoperability [7]

Metadata describes what the resource is, such as the subject keywords, how to use it, and how the resource is to be managed. We can say that metadata are tags that identify learning objects and educational resources with the possibility of being verified by a third party [8], ensuring accessibility to the descriptions of the objects and digital resources. Among the most common metadata standards for these purposes are IEEE LOM [9], Dublin Core Metadata Initiative [10] and SCORM [11]

The term interoperability refers to the ability to have two or more systems to exchange information and then, reuse that information [12]

While explaining why and how OER are labeled, we will get to know the system that contains them. Lynch [13], defines a repository as a computer system where multiple databases or files are located for distribution over the Internet. It is a data provider that integrates a set of services that incorporate, collect, preserve, consult and support management and dissemination of digital resources properly classified, to community members, through a Web interface.

Search engines are programs that track down documents according to specified keywords and return a result list with a brief description of the Websites or documents found related to search criteria. Currently, there are many engines using different software browsers, some are Web crawlers, also called bots or spiders that are designed to index Web pages and find words contained in those pages.

Meta-connector, a synonym of metasearch is a search engine, metadata aggregator, infomediary type, and infomediary is the term that results from the combination of the words: information and intermediary [14] A meta-connector is a Website that gathers and organizes large amounts of metadata and acts as an intermediary between those who need and those who provide information, as primary sources of information and repository providers. The metasearch can be set up to perform federated search or harvesting metadata via the Open Archive Initiative Protocol Metadata Harvesting (OAI –PMH) [15]

The harvest of metadata is a semi-automated process, which is led by a person with library and information systems training. It is a search on demand, in real time, it uses a central container that temporarily stores records from related repositories. Among its characteristics, it requires less search processing time; therefore, less response time.

We can say that the OAI-PMH is a low barrier mechanism for repository interoperability [15]

3. Contextual Framework

One of the main lines of work in the Open Education Movement, focuses on the production, spread, use and reuse of open educational resources. Currently, global agencies

such as UNESCO and Education For All (EFA), among others, promote projects aimed at the creation, use and processing of REA, and the development of repositories and systems that support and sustain its purpose, convinced that knowledge is a driver of economic development and growth in developing countries [16]

Haddad and Draxler [17], highlight that the digital content repositories also known as contentware, represents a crucial and challenging issue for organizations and educational institutions.

Thus, there were the initiative to develop an interagency basis developing a meta-connector that allowed access to the metadata of different digital educational repositories, that provides interoperability through the OAI-PMH [18]

As a basis for this project, there were the following meta-connectors:

- OA- Hermes. Mexican Metasearch engine, [19]
- Collector Open Science. Recolecta, [20]
- Diffusion in Red Alert is one of the largest Bibliographic free access portals, whose primary purpose is to give greater visibility to Hispanic literature, [21]
- Network of Scientific Journals from Latin and the Caribbean, Spain and Portugal [22]
- Open Access Platform for Electronic Scientific Journals Latin American Spanish Federated [23]
- Institutional Repositories for Scientific Documentation. Project with members from 8 countries, Argentina, Brazil, Chile, Colombia, Ecuador, Mexico, Peru, Venezuela [24]
- The Connecting Repositories CORE by JISC Meta-connector, [25]
- Digital Repository Infrastructure Vision for European Research, harvested regularly from more than 249 repositories, from 33 countries [26].
- Find the pearls. It is a catalog of millions of records representing open access resources, fueled by open access collections worldwide using OAI –PMH [27]
- Bielefeld Academic Search Engine. Allows searches of open academic resources. Operated by the Library of the University of Bielefeld offers OAI –PMH URL validation [28]
- ARIADNE Foundation European Partnership is open to the world to share and reuse knowledge. The core of the ARIADNE infrastructure is a distributed learning object repositories network [29].
- The Global Learning Objects Brokered Exchange - GLOBE was established from the following founding members: the ARIADNE Foundation in Europe, education.au in Australia, LORNET in Canada, the U.S. and MERLOT in NIME in Japan, [30]
- The OpenCourseWare via RSS feeds [31]
- Metasearcher via RSS feeds [32]
- JOCW. (2011). Japan OpenCourseWare Consortium, [33]

- Intute, comprising the Universities of Nottingham, Birmingham, Bristol, Heriot- Watt, Manchester Metropolitan and Oxford Merlot that exhibits REA refereed to academic consortium of different suppliers. The Intute site was closed on July 2011.

4. Working Strategy

The working strategy consisted in form works groups and collaborated remotely with the use of information and communications technology in videoconferencing sessions by Internet 2 and digital tools like blogs and discussion forums. A technical committee to know the metadata harvesting process was formed. A thesaurus metadata is also selected. Institutions that had a repository system (ITESM, ITCH and UdeG) contributed their repositories to relate to metaconector.

This development was part of Metaconector Educational Repositories project to promote the use of Learning Objects and Open Educational Resources: Best Practices, proposed and led by the Instituto Tecnológico de Estudios Superiores de Monterrey (ITESM), sponsored by the University Corporation for Internet Development (CUDI) and the National Council of Science and Technology (CONACYT), in the period from January to October 2011, with the participation of the University of Guadalajara (UdG), Montemorelos University (UM), Technological Institute of

Chihuahua (ITCH) and the Autonomous University of Guadalajara (UAG). As we can see, it was an interagency project.

The methodology consisted to integrate working groups and collaborate remotely with the use of information and communications technology in videoconferencing sessions by Internet 2 and digital tools like blogs and discussion forums. A technical committee to study the metadata harvesting process was formed. In that time, only three universities had an educational repository (ITESM, ITCH and UdG) and contributed with them to relate to meta-conector.

The roles involved to setting up the meta-search were: Moderator [M], person who coordinates the members of the team, Technician [T], responsible for the repository that will be related. Librarian [L], or person with related studies in information science and library, and a Programmer [P], responsible for developing the software application.

General Steps:

1. Analysis of digital repositories, data providers.
2. Define rules to import metadata per collection.
3. The search results are presented on the Website educonector.info
4. Monitoring, expect enter from another repository and update frequency of harvest

See on figure 1 the process followed to develop meta-conector.

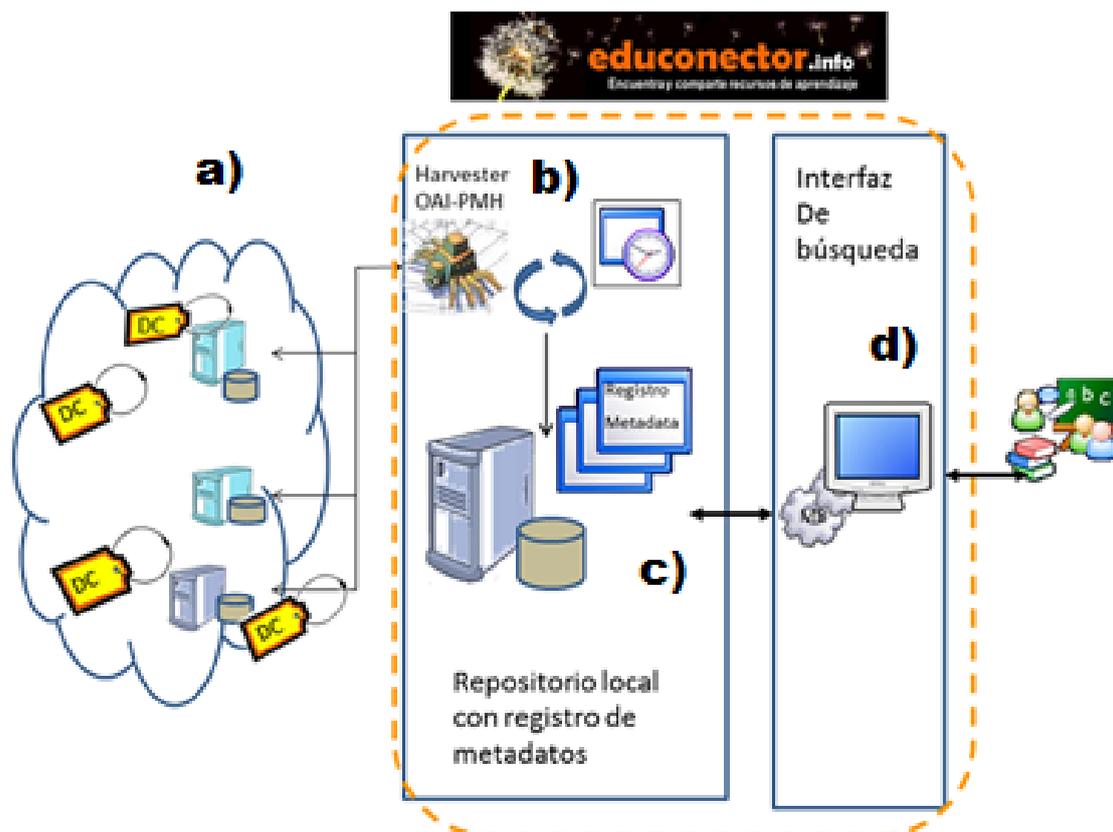


Figure 1. Diagram of the process development Source: Adame et al [14]. educonector.info observatorio abierto de producción académica y científica mexicana. Cap.3, p. 45.



Figure 2. Google Map, with the actual distribution of Mexican repositories related to educonector.info. Source: own authorship

5. Meta-connector Technical Specification

Educonector.info is the name of a metasearch engine that through the communication tool of Internet -based network, OAIconnect [15], allowed the linking of different Mexican digital repositories of open educational resources through metadata harvest intermediate. The metadata were interpreted considering the Dublin Core standard, and stored on a local server that serves as a repository of metadata while creating an index to facilitate the implementation of search mechanisms on a Web interface. See in figure 2, a) Repositories of OER interoperables by OAI-PMH b) Harvester, search engine c) Generating step of infomediary catalog d) Search interface.

The meta-connector educonector.info was setting up with general public licence (GPU) software, like the platform OAI. Connect and the learning management content (LMC) Drupal, system used to setting up dynamic websites. This software allows to publish, manage and organize big data content in a website.[14]

Most repositories related to educonector, are organized in collections, so the spider program of the educonector,

searches the OER metadata in the selected collections. Previously each repository is analyzed for a selective harvesting of it collections. With this step a repository profile is created for document the reliable information that allows to define harvesting metadata rules.

Few questions, used to design and to document the profile of each collection were:

What is the subject, discipline or knowledge area that the collection is specialized? What is the volume of records collection? What is the periodicity of harvest?

The OAI Protocol for Metadata Harvesting, provides a framework for interoperability of applications based on metadata harvesting. According to Lagoze et al [18], there are two types of actors within OAI- PMH:

- A) Data providers are repositories that expose structured metadata via OAI-PMH.
- B) Service providers then make OAI-PMH service requests to harvest that metadata-OAI- PMH is a set of six verbs that are used for metadata harvesters (service providers) to collect metadata; each verb has a unique purpose and meaning that make it easier to analyze the data. See table 1.

Table 1. Types of petition in the OAI-PMH.

OAI Request	OAI Response
Identify	Provides basic information repository as the repository name, base URL, protocol version, the first registration date, granularity, support deleted records, e- mail address of the repository.
ListSets	Provides a list of collections that have been established in the repository.
ListMetadataFormats	Provides a list of metadata format that are supported by the repository.
GetRecord	Provides a unique identifier to unambiguously identifies an item within a repository.
ListRecords	Facilitates the metadata for each record that meets the specified criteria.
ListIdentifiers Provides	Provides basic information for each record in the repository that meets the specified criteria.

Source: Adame et al [14]

6. Educational Impact and Results

- Through a simple web interface provide links to open educational qualified contents from Mexican Universities.
- At the beginning of the research project only three universities had an Open Educational Resources Repository, and in two years of operation, the meta-connector demonstrated an increment to visits to the three repositories as Pacheco and Cruz [34] related.
- Students and teachers that knew the meta connector, started to search for quality information in the three repositories, and the first impact comparing with open web like Google, was that there were just few resources, at 2013 the amount of OER per repository were DAR - 319, CREA - 191, ITCH - 180.
- According to users of meta-connector the results of search were better to find reliable content.
- These technological activities enabled the reflection on semantic web to describe digital educational resources through metadata content area.
- As a result of the educational impact and the technological infrastructure optimization, among other cooperation initiatives and the Open Educational Resource Movement, educonector info promotes the creation of new open repositories, with the opportunity to have Mexican presence in the world, through the ODOAR. On 2013 there were 11 repositories from 9 universities linked to the educonector. Actually educonector.info is down of web, by administrative policies. You can reach the repositories on RIMETRIC <http://azul.iing.mx1.uabc.mx/rimetric/>.
- Escuela de Graduados del Tecnológico de Monterrey (ITESM) or DAR: Develop, Learn and Reuse.
- Instituto Tecnológico de Chihuahua (ITCH) or Expovision or Portal SINED - Multimedia Podcast: National System of Distance Education.
- Universidad de Guadalajara (Ude G), CREA: Resource Center for Teaching and Learning.
- e- Gnosis: Digital Magazine Science and Technology.
- Institutional Repository of the Universidad Veracruzana .
- Universidad Autonoma of Nuevo León (UANL), UANL Institutional Repository.
- Universidad de las Americas, Puebla (UDLAP), CIRIA: Interactive Resource Center Information and Learning.

- Universidad Autonoma de Mexico (UNAM) Institucional Repository CUAED.
- Universidad Autonoma del Estado de Hidalgo (UAEH), Digital Library Repository.
- Instituto Tecnológico de Estudios Superiores de Occidente (ITESO - Guadalajara) Repository EduDoc - Education Documentation Centre.
- Universidad Autonoma del Estado de Hidalgo (UAEH) Digital Library.
- Instituto Tecnológico de Estudios Superiores de Occidente (ITESO-Guadalajara) EduDoc- Centro de Documentación sobre Educación.

7. Conclusions and Future Work

Based on the result of the develop of educonector.info, the aim of raising awareness of the open education movement, was achieved. First, among the twelve researchers of the participating universities. Later the students and teachers of their universities.

When we worked with the technical committee, we learned about the OAI-PMH protocol, the Dublin Core metadata standard, the structure of the repositories and search engines in addition to Creative Commons licensing.

The experience to work from different States of the Mexican Republic was one of the most significant challenges, it demonstrates that is possible if all the researchers are committed with the project. The use of Internet 2 was very useful for the remote meetings.

About the 9 repositories, related to meta-connector, 8 use a Dspace platform, and Dublin Core metadata.

Dublin Core facilitates the repositories harvesting, but is not the only metadata standard that we can use. The Simple Dublin Core standard, does not provide many of the attributes needed to tag mobile or multimedia OER, like photos or podcast; therefore, a possible future research work, will be about the inclusion of LOM metadata to repositories.

Is important to write a guidance for setting up repositories over GNU software, that allow the technicians save time and money in the development process.

If we compare the number of open repositories in USA and Europe, we can see that in Latin America we have a considerable lag, in areas such as e-Science, e-Journals and Open Access. Therefore it is necessary to disseminate, promote, research and innovate in this area. Over all, it is

necessary to develop the culture of producing reliable educational resources, to share under Creative Commons licensing.

We suggest considering the uniform development of institutional repositories. And motivate other Mexican and Latin American universities to participate in the open access movement, as well as think of a core metadata ad hoc to Spanish-speaking academic characteristics.

We need think about how to manage the increasing data and archive size of the new OER.

At the moment, educonector site is closed by administrative issues, the repositories can be reach on the site RIMETRIC available from <http://azul.iing.mx/uabc.mx>

In Mexico, we have a big potential to innovate the use of ICT and to produce OER that improve learning opportunities.

Acknowledgements

The authors of this study thank the National Research Council for Science and Technology (CONACYT) and to the University Corporation for Development of Internet (CUDI), for sponsoring this project. Also, members of the 4 universities who collaborated throughout the project.

REFERENCES

- [1] Van de Sompel, H. The future of the past of the Web. 3rd Joint Web archiving workshop, Online available from <http://www.jisc.ac.uk/events/2011/10/futureoftheWeb.aspx>
- [2] AMIPCI. Hábitos de los usuarios de Internet en Mexico. Asociación Mexicana de Internet, Online available from <http://www.amipci.org.mx/?P=editomultimediafile&Multimedia=348&Type=1>
- [3] Adame, S.I., Lloréns, L., Schorr, M. (2013). Retrospectiva de los repositorios de acceso abierto y tendencias en la socialización del conocimiento. *Revista Electrónica de Investigación Educativa*. REDIE Vol 15, No. 2
- [4] Budapest Open Access Initiative (2001). Budapest Initiative for Open Access., Online available from www.opensocietyfoundations.org/openaccess/translations/spanish-translation
- [5] Atkins, D., Brown, J. And Hammond, A. (2007). Report to The William and Flora Hewlett Foundation. A Review of the Open Educational Resources (OER) Movement: Achievements Challenges, and New Opportunities, Online available from <http://www.hewlett.org/uploads/files/ReviewoftheOERMovement.pdf>
- [6] Butcher, N., Ashka, K., And Stamenka, U. (2011). What Are Open Educational Resources. A Basic Guide to Open Educational Resources (OER). 1st Ed. Vancouver: Commonwealth of Learning. Pp. 5
- [7] Burgos, J. V. (2010). Knowledge sharing and free access to information for Open Educational Resources (OER). Digital magazine *educ@tion*, 143, Online available from http://www.educoas.org/portal/la_educacion_digital/143/articulos/reavladimirburgos.pdf
- [8] W3C (2010). World Wide Web Consortium, Online available from <http://www.w3.org/>
- [9] LOM (2010). LOM working draft v4. 1, Online available from <http://ltsc.ieee.org/doc/wg12/>
- [10] DCMI (2011). Dublin Core Metadata Initiative, Online available from <http://dublincore.org/>
- [11] SCORM (2010). Sharable Content Object Reference Model, Online available from <http://scorm.com/>
- [12] Castro, L., Lopez, G. (2012). Proposal for Measuring the Interoperability of Learning Objects. In M. Prieto, J. Melo and D. Pardíñaz, *Digital Resources for Education and Learning* (pp. 90-97). Merida, Mexico: Technological Institute of Mérida, Yucatán.
- [13] Lynch, C. A. (2003). Institutional repositories: Essential infrastructure for scholarship in the Digital Age. ARL Bimonthly Report, 226, 1-7, Online available from <http://www.arl.org/resources/pubs/br/br226/br226ir.shtml>
- [14] Adame, S.I., Burgos, V., Lloréns, L. (2013). *educonector.info* observatorio abierto de producción académica y científica mexicana. Cap.3, pp. 35-51. En Mortera, F. y Ramírez, M. S. (Coords). (2013). *Conexión de repositorios educativos digitales: educonector.info* Ed. LULU.com, México.
- [15] OAI (2012). Open Archive Initiative, Online available from www.oaiconnect.org/
- [16] UNESCO -EFA (2005). Monitoring Report for EFA in the World. Education for All – The Quality Imperative <http://www.unesco.org/new/es/education/themes/leading-the-international-agenda/efareport/reports/2005-quality/> accessed August 15, 2012.
- [17] Haddad, W., and Draxler, A. (2002). Technologies for education: potentials, parameters and prospects; Challenges and possibilities of ICT for education, UNESCO and the Academy for Educational Development, Online available from <http://unesdoc.unesco.org/images/0011/001191/119129e.pdf>
- [18] Lagoze, C. And Van de Sompel, H., The Open Archives Initiative Protocol for Metadata Harvesting. Protocol Version 2.0 of 2002-06-14, Online Available from <http://www.openarchives.org/OAI/2.0/openarchivesprotocol.htm>
- [19] Castro, A. And Garcia, E. (2007). OA -Hermes: academic metasearch, results of progress 2006-2007, National Autonomous University of Mexico, Online available from http://ict.udlap.mx/people/alfredo/rabid/reportes-final/unam/reportes_tecnico_unam.doc
- [20] Recolecta. (2011). Collector Open Science <http://www.recolecta.net/buscador/> accessed March, 2011.
- [21] Dialnet. (2011). Diffusion in Red Alert, Online available from <http://dialnet.unirioja.es/>
- [22] REDALYC. (2011). Network of Scientific Journals from

- Latin America and the Caribbean, Spain and Portugal
<http://redalyc.uaemex.mx/> accessed May 2011.
- [23] e-revistas. (2011). Electronic Scientific Journals, Online available from <http://www.erevistas.csic.es/>
- [24] Federated Network. (2011). Institutional Repositories for Scientific Documentation, <https://sites.google.com/site/bidclara/>
- [25] CORE. (2011). CORE by JISC, Online available from <http://core.kmi.open.ac.uk/>
- [26] DRIVER. (2011). Digital Repository Infrastructure Vision for European Research, Online available from <http://www.driver-community.eu/>
- [27] OAIster. (2011). Online available from <http://oaister.worldcat.org/>
- [28] BASE. (2011). Bielefeld Academic Search Engine, Online available from <http://oval.base-search.net/oval>
- [29] ARIADNE. (2011). Foundation, Online available from www.ariadne-eu.org
- [30] GLOBE. (2011). The Global Learning Objects Brokered Exchange – GLOBE, Online available from <http://www.globe-info.org/>
- [31] Meta-connector. (2011) OpenCourseWare, Online available from <http://www.ocwsearch.com/>
- [32] Universia. (2011). <http://ocw.universia.net/es/buscador.php>
- [33] JOCW. (2011). Japan OpenCourseWare Consortium, Online available from www.jocw.jp
- [34] Pacheco, A., Cruz, R. (2013). educonector.info observatorio abierto de producción académica y científica mexicana. Cap.5, pp. 68-80. En Mortera, F. y Ramírez, M. S. (Coords). (2013). *Implementación de algoritmos para la construcción de proveedores de datos OAI*. Ed. LULU.com, México.