A methodology for evaluating university environmental thought

Una metodología para evaluar el pensamiento ambiental universitario

DOMÍNGUEZ, Mercedes Monserrat†*, LINARES, Gladys, ESPINOSA, Araceli and HERNÁNDEZ, J. Santos

Benemérita Universidad Autónoma de Puebla, ICUAP, Postgraduate in Environmental Sciences, Mexico.

ID 1st Author: *Mercedes Monserrat, Domínguez* / **ORC ID:** 0000-0002-0559-0551, **CVU CONACYT ID:** 609864

ID 1st Coauthor: *Gladys, Linares* / **ORC ID:** 0000-0003-0856-3994, **CVU CONACYT ID:** 121612

ID 2nd Coauthor: *Araceli, Espinosa* / **ORC ID:** 0000-0003-0475-6932, **CVU CONACYT ID:** 48806

ID 3rd Coauthor: *J. Santos, Hernández* / **ORC ID:** 0000-0001-6286-8487, **CVU CONACYT ID:** 218016

Abstract

The socioeconomic and environmental attributes that characterize the environmental thought are the elements used to design a methodology to evaluate the substantive functions of teaching and research in the university environment, in a public higher education institution in Mexico. The proposed methodology is based on tools and instruments to generate knowledge of big data information, making use of: Information and Communication Technologies; of the principles of multivariate data management that consider the distinctive uncertainty of environmental phenomena; and the generation of multidisciplinary databases to apply data mining and obtain information with a systemic approach. The proposed tool with its mixed approach applies to the operation and multimedia text database study and from there the KDD process knowledge is acquired. The application context of the proposed methodology is the information in text form, published electronically, in the educational offer of the 2017 undergraduate and postgraduate admissions at the Benemérita Universidad Autónoma de Puebla (BUAP).

KDD Process, Data Mining, Environmental Education

Resumen

Los atributos socioeconómicos y ambientales que caracterizan al pensamiento ambiental son los elementos para diseñar una metodología de evaluación de las funciones sustantivas de docencia e investigación en el ámbito universitario, en una institución de educación superior pública de México. La metodología propuesta se basa en las herramientas e instrumentos para generar conocimiento de los datos de la big data, haciendo uso de: las Tecnologías de la Información y las Comunicaciones; de los principios del manejo multivariado de datos que consideran a la incertidumbre distintivo de los fenómenos ambientales; y de la generación de bases de datos multidisciplinarias para aplicar minería de datos y obtener información con un enfoque sistémico. La herramienta propuesta con su enfoque mixto se aplica en la operación y estudio de datos de texto multimedia y se obtienen conocimientos del proceso KDD. El contexto de aplicación de la metodología propuesta es la información en forma de texto, publicada vía electrónica, en la oferta educativa de licenciaturas y posgrados en el ingreso de 2017 en la Benemérita Universidad Autónoma de Puebla (BUAP).

Proceso KDD, Minería de Datos y Educación Ambiental

---

* Correspondence to Author (Email: monserrat.domingueznr@outlook.com)
† Researcher contributing first author.

Introduction

Environmental Sciences are defined as the set of sciences that identify, explain and investigate the relationships that the human being maintains as part of society, with himself and with his environment (Leff, 2006), that is, they consider in a whole, the elements of the society-nature binomial and their interactions.

Environmental thinking as part of the philosophy of Environmental Sciences is the result of the interaction of social and natural sciences, with an interdisciplinary and holistic vision of multiple causes and effects, knowledge and multi-interrelations (Leff, 2005).

Focusing these characteristics on university education, it highlights that Higher Education Institutions (HEIs) have been given the responsibility of forming a society that understands environmental problems and participates in generating knowledge and therefore integral solutions from environmental spheres, social and economic.

As María Teresa Bravo argues

"There are several aspects for which attention is paid to HEIs to address the environmental problem, among them: their potential as they are focused on science and the generation of technologies; its contribution to the solution and prevention of environmental degradation and its contribution to the construction of desirable development scenarios; due to its new mode of production of knowledge and skills, due to its systemic-complex nature" (Bravo-Mercado, 2012:17).

These particularities, added to the heterogeneous and constantly changing characteristics in the university environment (Eschenhagen, 2007), as well as the international, national or university association guidelines, frame environmental thinking in this field, and allow identifying the scarcity of tools to evaluate it.

To carry out the evaluation of environmental thinking in the university educational offer, in this work a tool is proposed to extract data, generate knowledge and identify patterns of multimedia data.

This methodology is aimed at exploring and identifying, and subsequently evaluating the data published in electronic media, applying the processes of handling and use of big data data.

In particular, the techniques of the process of "generation of knowledge from databases" or "knowledge discovery in data bases" (KDD) are used to text-type multimedia data, related to environmental thinking in the university educational offer.

The objective of this work is to design a methodology to delimit and subsequently evaluate the attributes of university environmental thought, at the higher levels in one of the most important macro-universities in Mexico, such as the Benemérita Universidad Autónoma de Puebla (BUAP), in a determined moment, through the application of the process of "generation of knowledge from databases" and the tools of text mining, considering uncertainty as a characteristic of the environmental vision.

This work is organized in four parts:

- In the first, the theory of environmental thought and the concepts of multimedia data are outlined; the process of generating knowledge of data management as part of the big data tools; and the principles of fuzzy logic.

- In the following section, the characteristics of environmental thinking in the educational offer of study at BUAP are delimited.

- The third part details the proposed methodology, based on multimedia mining and the KDD process.

- This methodology is applied to the educational offer of: 84 undergraduate degrees, 46 master's degrees and 22 doctorates, upon admission to BUAP in August 2017.

- Finally, the results generated from the designed methodology are shown and finally the conclusions are presented.

Theoretical framework

Currently, the interest in managing and obtaining knowledge and results from big data data is growing in multiple areas of knowledge, and Environmental Sciences are no exception.

This work is located from a systemic and mixed vision, so it becomes necessary to demarcate the theory on which the methodological proposal and its characteristics are based.

Environmental thinking

From the perspective of Environmental Sciences, environmental thinking has several definitions, the following being those that are related to the subject of this work:

- Fritjof Capra argues that environmental thinking is the "thought that occurs by expanding the systemic approach from the social sphere to the material world" (Capra, 2002: 22).

- Luz Arabany Ramírez, establishes that environmental thinking is the sum of the principles of "thoughts: complex, systemic and transdisciplinary" (Ramírez, 2015: 1) and since this definition treats environmental thinking with a holistic vision, it is immediately delve into each of these principles.

– Complex thinking is based on the arguments described and developed by Edgar Morin, defined by the author himself as one that:

– "It does not seek to unify everything separated in a thought of completeness, but to launch a challenge in pursuit of religion and complexity ... it is not a matter of approaching complexity as a pure rejection of the simplifying, of Aristotelian logic or of the principle of separation; on the contrary, complex thinking implies the integration of the above in a relational and continuous rotating principle" (Morin, 1992: 10).

– According to Kuhn: complexity is not an answer to the questions of the disciplinary sciences, but a challenge for thinking and the natural fusion of systems developments and information theory (Osorio, 2008).

– Systemic thinking is an approach to understanding reality, where the whole is more than the sum of its parts and with the study of the whole the parts and the connections between them are understood (Liévano and Londoño, 2012). It is identified in the multiple causes and multiple effects of the elements that make up a system and their relationship with the environment (Amozurrutia, 2006), in this case the socioeconomic and environmental perspectives.

– The transdisciplinary thinking proposed by Manfred Max-Neef, based on Morin's arguments, establishes that interdisciplinarity is the search for the unity of knowledge fragmented into disciplines, through the paradigmatic path (Osorio, 2012) and that the "Greening of thought" (Morin, 1991) "as a strategy for the social transformation of reality and for the constitution of the human condition" (Osorio, 2012: 244).

– Transdisciplinarity, as Nicolescu (2006) argues, goes beyond a new articulation of knowledge, a new conception of knowledge and a new understanding of reality must be considered as a new paradigm.

- Javier Gonzaga Valencia (2007) emphasizes that environmental thought "constitutes the structuring base of a society that respectfully and poetically inhabits the earth from the understanding of the complexity and permanent situations of chaos to which the complex system is subject also of life" (Valencia, 2007: 39).

Focusing on the field of university environmental education, it is considered that environmental thinking is not an imposition on HEIs, but a way of incorporating this paradigm into the current world, since "education is the strategy for achieving the objectives of the sustainability" (Gutiérrez and Martínez, 2010: 114).

The previous definitions of environmental thought allowed to identify the adjectives or attributes that characterize it, from the binomial that relates society to its environment.

Once the basic elements of the philosophy of university environmental thought have been cited, the following paragraphs outline the theory of big data on which the proposed methodology is based.

Multimedia and KDD data

To address the development of this article, the terms of big data are briefly presented, from general concepts to the processes to produce knowledge.

In our days, there is an enormous amount of data that is generated in real time and increases exponentially every day. The data can be public or private, and can be used and studied to generate information, knowledge or establish strategies for decision-making or marketing processes.

The large amount of data is known as big data. This term was coined in 2012 by Gartner, which defines it as "high-volume, high-speed and / or wide-variety information that requires new forms of processing to allow better decision-making, information discovery, and process optimization" (Bello -Orgaz, Jung, Camacho, 2015: 45).

This large volume of data is explained a few years later by Jenkins, Slomczynski and Dubrow (2016), who affirm that big data is made up of various sources; and Stone (2017) clarifies that, in the data, it must be considered: the frequency of emergence, the iterations that could occur between them and the constant updating of the model they represent to convert them into information. In addition, the large amount of data comes from heterogeneous sources and simultaneously (Joyanes-Aguilar, 2016) and can be manipulated thanks to modern information technologies and computational facilities.

The veracity of the results provided by said data depends on the model that represents them, the techniques for their interpretation (Bonomo, 2018), their quality and privacy.

Big data data, due to its organization and arrangement, are classified into data: structured, semi-structured and unstructured (Camargo-Vega, Camargo-Ortega, and Joyanes-Aguilar, 2015). In this work we have multimedia data, not structured in databases and without labels to identify them.

Multimedia data can be of four types: images, text, audio and videos.

To obtain information from these multimedia data, the analytical techniques of multimedia mining are used (Contreras-Barrera, 2014), which are a set of methodologies and platforms to analyze data that are not organized and that generate non-trivial information, that can be applied directly for: the search for knowledge, decision-making and the use, disclosure and publication of information (idem).

One of these tools is the one applied in this work, which is the KDD process "knowledge discovery in data bases" (by its name in English), defined as an iterative and interactive process that explores very large volumes of text data, exhaustively to determine relationships (Minerva Data Mining, 2018) and "marks a paradigm shift in which what is important is useful knowledge" (Riquelme, Ruíz and Gilbert, 2006).

The KDD process consists of five steps which are: selection, processing, transformation, data mining, and data interpretation and evaluation (idem).

Within KDD, data mining allows the extraction of implicit, previously unknown and potentially useful information from said data (Frawley, Piatetsky-Shapiro and Matheus, 1992) and is considered the step for interdisciplinary analysis processing (Rojas, Chavarro and Moreno, 2008).

As such, data mining is a tool for exploring the vast amount of unorganized data; establish models and / or patterns; find relationships or associations; and extract knowledge when instruments are applied to identify and interpret: trends, deviations (Montes-y-Gómez, 2001), groupings, summaries and sequence analysis, among many others.

Within data mining, this work only considers text mining, in particular the bag of words and its analysis techniques.

Ian H. Witten, Frank Eibe and A. Mark (2011) define text mining as a technique that searches, finds and explores interesting patterns in the text, that is, finds relevant information for a purpose, from texts in natural language. Ronen Feldman, and James Sanger (2007) add that it is a new "knowledge-intensive process in which a user interacts with a collection of documents over time, using a set of analysis tools" (Feldman and Sanger 2007: 1) to obtain knowledge of a set of texts from the "great amount of textual information" (Montes-y-Gómez, 2001: 1).

Among the most used text mining processes, there are: natural language processing, discourse analysis, content analysis, univariate and multivariate statistical methods, modeling and mathematical tools, and information retrieval and its data classification and grouping methods (Contreras-Barrera, 2014).

In this work, content analysis and multivariate statistics are used for data grouping of university environmental thought.

To obtain knowledge of the data in the form of texts, whether electronic or printed, Oviedo and his co-authors (2017) propose a tool, which first consists of carrying out a predictive analysis, where the texts are studied and transformed into some type of structured representation or semi-structured that will facilitate its subsequent study (Oviedo, Oviedo and Vélez, 2017); and then a descriptive analysis of the data to obtain useful information.

The series of steps in the methodology of Oviedo and his co-authors coincide with the steps of the KDD process for text processing, which are:

– Step 1, determine the purpose of the study.

– Step 2, retrieve the information by gather, identifying, collecting and validating the texts in the appropriate format for their study; which in KDD corresponds to the selection and processing stages.

– Step 3, perform the text processing, in order to define the keywords to study and eliminate unnecessary information; which in KDD is identified as transformation.

– Step 4, execute the extraction and analysis of classes, relationships, associations or sequences, in order to find evidence of concepts and existing structures, through the number of repetitions of the word in the document; identified within KDD as data mining.

– Step 5, present the results obtained (spreadsheets, databases, clusters, summaries, text markings, relationships, etc.) for their interpretation; which within KDD was defined as the last stage of interpretation and evaluation.

By applying the steps proposed by Oviedo and his co-authors (2017), within the framework of the KDD process, the "clustering" will be generated, where the data is represented in groups (Linares, 2006), "each data belongs only to one cluster and the set of them includes the total data" (Rojas, et. al., 2008: 2). In this work the cluster quantity was defined based on the three dimensions: environmental, social and economic of sustainability (idem).

Fundamentals of fuzzy logic

Finally, related to the theoretical framework on which the work is based, it is identified that uncertainty is a characteristic of environmental phenomena, since there is no aspect of Environmental Sciences where there is no uncertainty, as it would not be a real problem.

To include uncertainty in the evaluation of the attributes of university environmental thought, the principles of fuzzy logic are applied.

Fuzzy logic is identified by producing exact results from imprecise data (Morales-Luna, 2002), in a similar way to human thought (Jang, Mizutani and Sun, 1997) and "describes reality through flexible models that include the subjectivity and uncertainty" (Lazzari, Machado and Pérez, 2000: 6).

The fuzzy term, as cited by Flores and Camarena (2013) is an adjective of logic, which "refers to the degrees of significance or values in the measurement of the uncertainty of linguistic variables" (Flores and Camarena, 2013: 12).

This logic is multivalued (Morales-Luna, 2002) where the fuzzy is determined as the possibility of assigning different truth values to the variables, and each element of the universe is associated or corresponds (function) to a degree of belonging to the set in the interval [0,1] (Flores and Camarena, 2013) identified as a fuzzy set.

The fuzzy logic generates that each variable, in this case the frequency of characteristic attributes of environmental thinking in the educational offer of the BUAP, have a range of possible solutions, with interrelationships among themselves, and thus be an instrument to give value to the data obtained (Zadeh, 1965) through a mathematically represented fuzzy set (Flores and Camarena, 2013).

The fuzzy logic tools allow the incorporation of uncertainty as part of the proposed methodology in the attributes of environmental thinking.

Environmental thought at the Benemerita Universidad Autónoma de Puebla (BUAP)

Once the main theoretical issues of this methodological proposal have been cited, it is important to define the context in which this theory is applied.

Said context of study is the Benemérita Universidad Autónoma de Puebla, a public institution of upper secondary and higher education, leader in the central-eastern region of Mexico, formed by 43 academic units, distributed in the state of Puebla (Gaceta Universitaria BUAP, 2017).

The BUAP to carry out its substantive institutional functions of: teaching, research, extension and dissemination of culture (BUAP, 2007) is based on its institutional philosophy.

Within the institutional philosophy, the vision of environmental thinking is identified in the principle that it is a sustainable university that focuses on building desirable development scenarios and participating with society for its transition to sustainable lifestyles (BUAP, 2014).

In general, as María Luisa Eschenhagen and Enrique Leff argue, a university with a sustainable vision guides its organization and formulates the educational offer from the perspective of interdisciplinarity, as "a development potential that demands their right to citizenship and a passport to move freely by the traditional frontiers of knowledge" (Eschenhagen, 2008: 127).

From this generality, university environmental education due to its systemic-complex nature" (Bravo-Mercado, 2012: 17) has multiple definitions. In this work the definition of Alfredo Ávila-Galarza is used, understood as "the process of acquisition of values and clarification of concepts whose objective is to develop attitudes and capacities necessary to understand and appreciate the interrelationships between man, his culture and his biophysical environment" (Ávila-Galarza, 2003: 45) which is one of the theoretical pillars for this work.

In BUAP this integrative vision is identified, as it forms part of its institutional strategic planning; through the different development plans of its rectors in the last twenty years; and the implementation of their educational models.

The substantive functions in BUAP from the environmental thinking approach have been oriented since 2007 to three major actions, which are (BUAP, 2007):

– The training of professionals committed to sustainable development and therefore to the environment.

– Conducting multi and interdisciplinary research to propose integral solutions to global problems.

– Establish ties for the dissemination and connection of the institution with its environment and its proposals to specific environmental problems with holistic solutions.

Regarding teaching, which is the study area of this work, BUAP is guided by the educational and organizational model established in 2007, known as the Minerva University Model, which includes: the philosophy of social responsibility and environmental thinking as part of its strategic planning (BUAP, 2014).

Said model contains environmental parameters in the undergraduate study plans; and in the teaching of undergraduate and postgraduate degrees related to Environmental Sciences, with a multidisciplinary vision.

The principles of the philosophy of environmental thought at BUAP were based on the guidelines of events that had occurred in various geographic and administrative coverage, related to environmental education, both in Mexico and internationally.

The main guiding events were: at the international level, the guidelines and public policies of the United Nations Environment Program (UNEP) and those of UNESCO, both within the United Nations Organization (Bravo-Mercado, 2012) ; in Latin America, the characteristic approaches of respect for nature (provider, councilor and mother) (Leff, 2009) that include the knowledge and customs of the peoples who have inhabited these areas for centuries. Regarding the guidelines at the national level, national educational models and plans have been in place for several decades, mainly those of ANUIES (Olaguez and Piero, 2013); in addition to the proposals generated by the meetings of Higher Education Institutions (COMPLEXUS, 2013), and the guidelines arising in forums, academies, collaboration networks or results of executive assemblies to propose joint and multidisciplinary work.

The study of these guidelines, focused on institutional substantive functions, were the basis for obtaining the attributes that characterize university environmental thought.

These attributes were obtained by applying text mining software to the guidelines described above and with this, the socioeconomic and environmental elements that determine university environmental thought were identified (procedure that is detailed in the following section), in addition to adding the results of previous research (Domínguez, 2015) in relation to the preferences that the university community expressed on the subject under study.

The delimitation of the attributes that characterize environmental thought was the first step of the proposed evaluation methodology, which is described in the following paragraphs.

Methodology for the study of environmental thinking in the educational offer of admissions 2017

To evaluate the principles of university environmental thought, a mixed methodology was designed, which applies brainstorming techniques, consultation with experts and content analysis within the qualitative methodology; as well as the tools of the descriptive process of generating knowledge of databases, in its various stages for quantitative analysis.

This methodology was structured in two stages: in the first, the attributes that characterize the philosophy of university environmental thought were specified (Eschenhagen, 2008) and, in the second, based on the study design of Oviedo and his co-authors (2017) the tools of the knowledge acquisition process in databases were applied to classify texts and generate groupings and associations (Montes-y-Gómez, 2001) in written documents (printed and / or electronic) in the studied context.

The steps of the first stage were:

- The environmental education guidelines were analyzed in different coverages as already mentioned in the previous section, both from meetings at the international level, proposals in the Latin American and national scope and the agreements of meetings of directors or academics, giving place to proposing a series of adjectives (attributes) that characterize university environmental thought.

The analysis was carried out by applying the electronic tools of the text mining software to obtain the words and phrases that are most often cited in the texts of these guidelines.

The result was the identification of the 65 attributes, which characterize university environmental thought and are: development, tenable development, sustainable development, sustainability, tenability, biodiversity, climate change, environmental sciences, complexity, environmental pollution, basins, environmental deterioration, dimension ecological, ecology, ecosystems, ecotechnics, environmental education, systemic approach, ecological balance, environmental impact, environmental indicators, interdiscipline, environmental research, environment, environmental methodology, multidisciplinary, environmental paradigm, environmental thinking, environmental problems, natural resources, environmental risk , environmental services, engineering techniques applied to the environment, systems theory, transversality, alternative developments, economic dimension, economic ecology, ecological economy, globalization, environmental anthropology, environmental architecture, environmental culture, social dimension, sociocultural dimension, ecodevelopment, education, ethnography, environmental ethics, environmental history, multiculturalism, social responsibility, environmental knowledge, socioeconomic, environmental sociology; international agreements, environmental law, institutional dimension, environmental guidelines, environmental management, environmental governance, environmental institution, environmental planning, environmental policy and environmental programs.

-    In addition to these results, a brainstorming was carried out with a group of professionals in Environmental Sciences in order to demarcate said attributes and point out that they are characteristic of environmental thinking and include its socioeconomic and environmental aspects.

-    As a requirement of the software used, it was required to group the 65 attributes into categories for later analysis.

With a holistic intent, 9 categories were proposed, which are: sustainability, environment, interdiscipline, ecotechnics, transversality, globalization, social responsibility, environmental policy and complexity.

With the characteristic attributes of the nature-society binomial, university environmental thought and the 9 categories of analysis, a consultation was carried out with 12 experts (Briones, 2001) on the subject, all of them currently working with the principles of environmental epistemology.

The consultation had the purpose of knowing, based on experience, in which of the 9 suggested analysis categories the attributes of university environmental thought already identified are located. Subsequently, these criteria were grouped through "mode" (statistical concept of central tendency).

The results obtained were the following:

–   Sustainability, grouped 6 attributes that are: development, tenable development, sustainable development, tenability, sustainability and eco-development.

–   The environment consisted of 16 attributes: biodiversity, climate change, environmental pollution, basins, environmental deterioration, ecological dimension, ecology, ecosystems, ecological balance, environmental impact, environmental indicators, environment, environmental problems, natural resources, environmental risk and environmental services.

–   Interdisciplinary, included 6 attributes: interdisciplinary, environmental research, environmental methodology, multidisciplinary, environmental paradigm and environmental thinking.

–   Eco-techniques, encompasses the attributes of engineering techniques applied to the environment and eco-techniques.

–   Transversality only has the attribute of transversality.

–   Globalization, is formed by 4 attributes that are economic dimension, economic ecology, ecological economy and globalization.

– Social responsibility, was integrated by 16 attributes, which are: alternative developments, environmental anthropology, environmental architecture, environmental culture, social dimension, sociocultural dimension, education, ethnography, environmental ethics, environmental history, multiculturalism, social responsibility, environmental knowledge, socioeconomic, environmental education and environmental sociology.

– Environmental policy, grouped into 10 attributes, which are: international agreements, environmental law, institutional dimension, environmental guidelines, environmental management, environmental governance, environmental institution, environmental planning, environmental policy and environmental programs.

– Complexity, with 4 attributes that are: environmental sciences, complexity, systemic approach and systems theory.

With the 65 attributes and their grouping into analysis categories, the first stage of the proposed methodology was completed.

In the second stage, the tools to generate knowledge of big data are applied, particularly with the KDD process and the five steps proposed by Oviedo and his co-authors (2017), to evaluate the attributes determined in the first stage and thereby generate results. These steps are:

- Step 1, information published electronically on the official page of the BUAP institution (www.buap.mx) of the educational offer of admission in 2017 was identified.

- Step 2, the data of the educational entry offer in August 2017 were selected and retrieved, having as a universe 152 study programs.

- Step 3, the published study programs were cleaned, discarding the information not related to the topic.

- Step 4, content analysis techniques were applied, defined by Andréu as "a technique for interpreting texts, whether written, engraved, painted, filmed ..., or in a different way where all kinds of data records can exist, the common denominator of all these materials is their ability to host content that, read and interpreted properly, opens the doors to the knowledge of various aspects and phenomena of social life" (Andréu, 2000: 2).

This content analysis was done with Atlas Ti, which is a software tool for the analysis of qualitative data.

Atlas Ti is characterized by the handling of large volumes of textual data, which organizes, regroups and manages them based on the definition of "analysis codes" and thus transforms the analyzed data into information in spreadsheets or databases for its subsequent interpretation and descriptive analysis (Muhr, 1997).

Using the Atlas Ti tools, the presence of the 65 attributes was detected in the 152 programs of the educational offer, resulting in the number of repetitions of each of the attributes in the published documents of the educational offer of degrees and postgraduate students upon entering BUAP in 2017 as already explained.

Atlas Ti generated a spreadsheet that groups the frequencies of the identified attributes.

The attribute frequency database was transformed into category frequency, based on the classification already defined in the first stage of the study, obtaining the maximum number of frequencies in each category.

- Step 5, the category frequency spreadsheet is the basis for the quantitative analysis of data and the generation of knowledge from them.

Based on this information, two types of analysis were carried out: the study of the data with the tools of descriptive statistics and the application of the text mining instruments within the KDD process.

Regarding the descriptive statistics, the analysis by areas was carried out to identify the percentage of citations of the attributes in the texts of the educational offer, based on the areas of sustainability and the nature-society binomial. These data and their results are presented in the next section.

For the text mining analysis, the maximum number of citations of each category was identified and a figure was calculated, corresponding to the quotient of the frequency of citations of each undergraduate and postgraduate degree between said maximum, generating a new spreadsheet, with the measurement of each category in fraction between [0 and 1] that includes the uncertainty in the characteristics and evaluation of the fuzzy logic.

The new spreadsheet was analyzed with the Minitab software (Minitab, 2010) using the conglomeration technique, which is a multivariate statistical technique, which starts from measuring the proximity between individuals to later search for the groups of individuals most similar to each other, according to a series of variables.

The purpose of this tool is to divide a set of objects into groups or clusters in such a way that the profiles of the objects in the same group are similar, that is, it creates groups or conglomerates with similar characteristics until a single block is obtained (Linares, 2006).

In this study, the hierarchical clustering technique was used, since it allows the variables to be agglomerated, starting from the calculation of the distance matrix and selecting from a wide variety of distance measures and grouping methods (idem). Distance measures evaluate the degree of similarity or proximity between two elements, in this study the distance correlation coefficient was used.

Regarding the linkage, the Ward link was selected, due to its characteristics that the loss of information when joining two elements is minimal, because it is the sum of squares of intra-cluster distance (idem).

The result obtained was a dendrogram, which is a type of graphical representation in the form of a tree, which summarizes the process of grouping the clustering technique of the elements of university environmental thought.

The results of this methodology are presented and analyzed in the following section.

Results

The present work was carried out with the information in texts published electronically on the educational offer at the higher level of: 84 bachelor's degrees, 46 master's degrees and 22 doctorates, published on the BUAP website (www.buap.mx) in the admision of 2017.

From the content analysis of the data of the 152 textual files examined, 1,188 citations of attributes for undergraduate degrees and 853 for postgraduate were obtained, in total 2,041 citations of the attributes that characterize environmental thinking.

The two tools already described above were applied to these data: descriptive statistics and later the study from text mining.

The results obtained were the following:

- The descriptive analysis is presented in Figure 1:

It is observed that 29.7% of the citations are related to the attributes grouped in the category of social responsibility; the next category for the number of citations was complexity, which accumulated 14.8%; environment follows with 14.%; in interdiscipline, 11.6% were identified; the concepts of sustainability presented 10.5% of the citations; Thus, we have ecotechnics with 7.6%, the attributes grouped in globalization were 3.9% and environmental policy was only identified in 3.5% of the total documents analyzed.
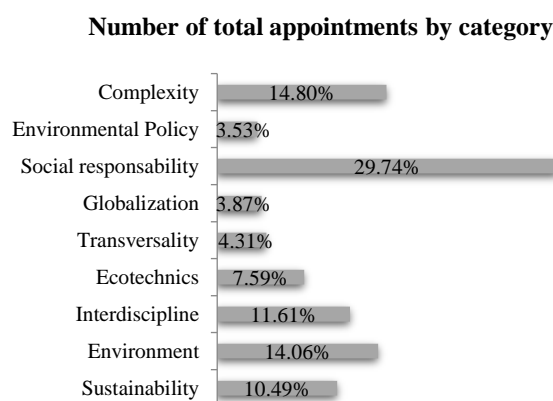
Number of total appointments by category

Category	Percentage
Complexity	14.80%
Environmental Policy	3.53%
Social responsability	29.74%
Globalization	3.87%
Transversality	4.31%
Ecotechnics	7.59%
Interdiscipline	11.61%
Environment	14.06%
Sustainability	10.49%

Figure 1 Total number of citations by category of analysis

From this descriptive analysis, it is observed that the philosophy of social responsibility is a topic that has been given greater importance within the institution and that the theory of complex systems has been considered as guidelines within the educational offer, given the number of citations that were identified in the texts studied.

These results suggest that the elements of environmental thought have been considered in the educational offer of higher education at BUAP, focusing on the issue of the institution's social responsibility as part of its substantive functions.

Transversality as a property of complex systems that characterize environmental thinking is only included in undergraduate degrees and therefore the number of citations.

- The text mining analysis, considering the fuzzy data and the application of multivariate clustering techniques, are presented in Figure 2.

- It was observed that the categories are grouped into the environmental and socioeconomic dimensions that characterize environmental thinking.
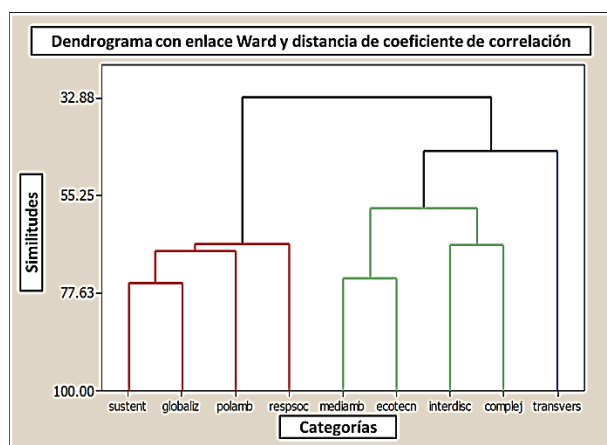

Figure 2 Dendrogram of similarity between analysis categories

- The grouping was in three "clusters": the first includes the categories of sustainability, globalization, social responsibility and environmental policy, which are identified with the socioeconomic area; the second group is made up of the categories of: environment, interdiscipline, ecotechnics and complexity that characterize environmental philosophy; and, finally, the third group, which is only the transversality that has been applied to the environmental education system.

- The categories that present the greatest similarity are: in the first cluster the categories of sustainability and globalization with 75.26%; and in the second group the categories of environment and ecotechnics, similar in 74.19%, since both are related in ecological aspects.

- The environmental categories present a similarity between the four, of 58.16% and together they are similar with transversality in 45.19%.

- Analyzing the educational offer as a system and applying the proposed methodology that includes uncertainty, it is seen that the three clusters are similar in 32.87%, that is, that the adjectives that characterize environmental thinking in BUAP are similar in equivalent proportions.

In short, the number of times the attributes were identified and their grouping into categories considering uncertainty, showed homogeneous groups, unknown a priori, which allowed to verify their integration and relationship with the environmental and socioeconomic areas of university environmental thought, within of the educational offer published in the admissions of 2017.

It should be noted that this methodological process of generating knowledge from databases can be used in texts (printed or electronic) related to environmental thinking in HEIs.

Conclusions

The analysis and the results obtained from the two stages of the methodological proposal allow to raise the following conclusions:

- The result of this work is evidence that the analysis of the phenomena of Environmental Sciences requires tools for the perception, interpretation and modeling of data, with a holistic approach and considering uncertainty as part of them, being fuzzy logic a tool for its study.

- The study of the principles of environmental thought, from the perspective of the environmental and the socioeconomic, offers a new way of incorporating the philosophy of environmental thought in the university environment into the interior of BUAP.

- The methodology proposed in this work, based on the tools of text mining and the KDD process, can be applied in any subject that has historical textual data that is available in electronic or physical media, related to university environmental thought, such as: educational offer, institutional plans, educational model, disciplinary, multidisciplinary or transversal subjects, research, thesis, etc.

References

Andréu, A. J. (2000). *Las técnicas de análisis de contenido: una revisión actualizada*, Fundación Centro Estudios Andaluces, Universidad de Granada, v.10, n. 2, p. 1-34, Consultation: 20/09/2016. Available in: <http://public.centrodeestudiosandaluces.es/pdfs/S200103.pdf

Amozurrutia, J. (2006). *Lógica Difusa y Redes Neuronales aplicadas a las Ciencias Sociales: un reto a la práctica interdisciplinaria*, Jornadas anuales de investigación 2005. México: CEIICH, UNAM.

Ávila Galarza, A. (2003). *La educación ambiental a nivel superior*. Rev. Teorema Ambiental. No. 37. Dec-Jan 2003. Available in: http://abiunsa.edu.pe/wp-content/uploads/2014/04/LA-EDUCACION-AMBIENTAL-A-NIVEL-SUPERIOR.pdf

Bello-Orgaz, G., Jung, J., y Camacho, D. (2015). *Social Big Data: Recent achievements and new challenge*, Information Fusion, num. 28, pp. 45-59.

Benemérita Universidad Autónoma de Puebla. (2007) *"Modelo Universitario Minerva: Fundamentos, Modelo Educativo Académico, Estructura Curricular, Integración Social y Documentos de integración"* BUAP, enero 2007.

Benemérita Universidad Autónoma de Puebla. (2014) *Plan de Desarrollo Institucional 2013 - 2017*, México.

Benemérita Universidad Autónoma de Puebla. (2017*) Gaceta Universitaria BUAP*, año XXXV, No. 211, febrero 2017.

Bonomo, H. (2018). *Las profecías autocumplidas de la tecnología.* Diálogo Político Año xxxiv, n.º 2 2018 Editor Konrad-Adenauer-Stiftung (Fundación Konrad Adenauer).

Bravo-Mercado, M.T. (2012). Coordinadora, *Los Planes Ambientales en la Educación Superior en México. Construyendo sentidos de sustentabilidad (2002-2007),* SEMARNAT – UNAM – INE y ANUIES, México

Briones, G. (2001). *Métodos y Técnicas de Investigación para las Ciencias Sociales*, México, Trillas.

Camargo-Vega, J. J. Camargo-Ortega, J. F. & Joyanes-Aguilar, L. (2015). *Conociendo Big Data*. Revista Facultad de Ingeniería, 24(38), 63-77. Retrieved May 08, 2019, from http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0121-11292015000100006&lng=en&tlng=es

Capra, F. (2002). *Las Conexiones Ocultas, Implicaciones Sociales, medioambientales, Económicas y Biológicas de una Nueva Visión del Mundo*. Barcelona: Editorial Anagrama

Contreras-Barrera, M. (2014). *Text mining: a current view*. Biblioteca Universitaria México, 17(2), 129-138

COMPLEXUS Consorcio Mexicano de Programas Ambientales Universitarios para el Desarrollo Sustentable. (2013). *Indicadores para medir la contribución de las Instituciones de Educación Superior a la Sustentabilidad.* Guanajuato México. Comité 2013-2015

Domínguez N. M. (2015) *Gestión Ambiental Sustentable para la Ciudad Universitaria de la Benemérita Universidad Autónoma de Puebla,* Tesis de Maestría en Ciencias Ambientales. ICUAP-BUAP

Eschenhagen, M. L (2007) *La educación ambiental superior en América Latina: una evaluación de la oferta de posgrados ambientales*. En: Revista Theomai. Available in: <http://revista-theomai.unq.edu.ar/Numero16/Eschenhagen. pdf

Eschenhagen, M. L (2008) *Retos epistemológicos y teóricos para el curriculum ambiental de posgrados,* Avaliação (Campinas) vol.13 no.1 Sorocaba, marzo 2008, versión impresa ISSN 1414-4077 versión en línea ISSN 1982-5765

Feldman, R., y J. Sanger, J., (2007) *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*, Cambridge: Cambridge University Press

Flores P.L. y Camarena L.M. (2013). *Evaluación de programas públicos en el marco de la realidad social, metodología basada en la lógica difusa como instrumento para el análisis de fenómenos sociales*. Revista Latinoamericana de Metodología de la Investigación Social - ReLMIS. N° 5. Año 3. Abril- septiembre 2013. Argentina. Estudios Sociológicos Editora. ISSN 1853-6190. Pp. 8 - 23. Available in: http://www.relmis.com.ar/ojs/index.php/relmis/article/view/58.

Frawley, W. J., Piatetsky-Shapiro, G., & Matheus, C. J. (1992). *Knowledge Discovery in Databases: An Overview*. AI Magazine, *13*(3), 57. https://doi.org/10.1609/aimag.v13i3.1011

Gutiérrez B. B y Martínez R.C., (2010) *El plan de acción para el desarrollo sustentable en las Instituciones de Educación Superior. Escenarios posibles,* Revista de la Educación Superior, Vol. XXXIX (2), No. 154, abril-junio de 2010, pp. 111-132. ISSN: 0185-2760

Jang, J., Mizutani, E. y Sun, C. (1997). *Neuro-fuzzy and soft computing: A computional approach to learning and machine intelligence*. New York: Prentice Hall. KOSKO, B. (1995).

Jenkins, J., Slomczynski, K., y Dubrow, J. (2016). *Political Behavior and Big Data*, International Journal of Sociology, n. ° 46, pp. 1-7

Joyanes-Aguilar, L. (2016). *Big Data, Análisis de grandes volúmenes de datos en organizaciones*, Alfaomega Grupo Editorial

Lazzari, L., Machado E. y Pérez R. (2000). *Los conjuntos borrosos una introducción*. Facultad de Ciencias Económicas, Universidad de Buenos Aires, Cuaderno No 9

Leff, E. (2005) *Complejidad, racionalidad ambiental y diálogo de saberes*, Ponencia presentada en el ciclo de conferencias "Los retos del siglo XXI y las ciencias sociales", del 27 de junio al 1 de julio de 2005, en el marco del 75 Aniversario del Instituto de Investigaciones Sociales, UNAM; México, Retrieved from http://conceptos.sociales.unam.mx/conceptos_final/470trabajo.pdf

Leff, E (2006), *Aventuras de la epistemología ambiental. De la articulación de las ciencias al diálogo de saberes*. México: Siglo XXI

Leff, E. (2009). *Pensamiento Ambiental Latinoamericano: Patrimonio de un Saber para la Sustentabilidad.* VI Congreso Iberoamericano de Educación Ambiental. San Clemente de Tuyú Argentina,   Retrieved   from http://www.cep.unt.edu/papers/leff-span.pdf

Liévano M. F. y Londoño, J.E. (2012). *El pensamiento sistémico como herramienta metodológica para la resolución de problemas,* Revista Soluciones de Postgrado EIA, Número 8. p. 43-65. Medellín, enero-junio de 2012

Linares F. G. (2006). *Análisis de Datos Multivariados*, BUAP Facultad de Ciencias de la Computación, Puebla, México.

Minerva Data Mining. (2018). *Descubrimiento de Conocimiento en Bases de Datos KDD Guía*, consultado en línea en https://mnrva.io/kdd-platform.html el 1 de mayo de 2019.

Minitab 17 Statistical Software (2010). [Computer software]. State College, PA: Minitab, Inc. (www.minitab.com)

Montes-y-Gómez, M. (2001). *Minería de texto: Un nuevo reto computacional*, Laboratorio de Lenguaje Natural, Centro de Investigación en Computación, Instituto Politécnico Nacional. accessed August 28, 2018 from https://ccc. inaoep. mx/~mmontesg/publicaciones/2001/MineriaTexto-md01. Pdf, el 28 de agosto de 2018

Morales-Luna, G. (2002). *Introducción a la Lógica Difusa*, Centro de Investigación y Estudios Avanzados del IPN, México.

Morin, E. (1991) *El pensamiento ecologizado* En: Gazeta de Antropología, Nº 12. Available in: http://www.ugr.es/~pwlac/G12_01Edgar_Morin. html

Morin, E. (1992). *El Método IV*: *Las ideas. Su hábitat, su vida, sus costumbres, su organización*. Madrid: Ediciones Cátedra

Muhr, Th. (1997). *Atlas/ti. Visual Qualitative Data. Analysis, management and model building*, User´s Manual and reference, Berlin, Scientific Software Development. (New version 4.1 design for Windows 95 and Windows NT).

Nicolescu, B. (2006). *Transdisciplinariedad: pasado, presente y futuro*. First part. In: Revista Visión docente con-ciencia, (31):15- 31. Available in: http://www.ceuarkos.com/Vision_docente/revista31/t3.htm. Second part. In: Revista Visión Docente Con-Ciencia, (32):14-33. Available in: http://www.ceuarkos. com/Vision_docente/revista32/t4.htmVisión

Docente Con-Ciencia Año VI, No 31, Julio-Agosto 2006

Olaguez T.E. y Piero, R. (2013). *Perspectivas de educación ambiental en las instituciones de educación superior.* Revista Iberoamericana de las Ciencias Biológicas y Agropecuarias: CIBA. Vol. 2, Núm. 3. Pp. 1-8. Centro de Estudios e Investigación para el Desarrollo Docente AC. Available in: file:///C:/Users/CA-BUAP-165%20RCAyDS/Downloads/Dialnet-PerspectivasDeEducacionAmbientalEnLasInstituciones-5063619%20(2).pdf

Osorio, G. S. N. (2008) *La emergencia de una nueva racionalidad*, En: Bioética y pensamiento complejo I: Un puente en construcción. Bogotá: Prontoprinter Ltda-UMNG: 19-55.

Osorio, G. S. N. (2012). *Ciencias de la Complejidad, Pensamiento Complejo y Conocimiento Transdisciplinar*. Re-pensando la Humana Coditio en un mundo tecnocientífico. La Crisis axiológica raíz de todas las crisis que sufre nuestro mundo: Cómo manejarnos con ella. En: España *ISBN:* 978-84-686-3345-9 *ed:* Cetr-Bubok Publishing S.L, *v.*, p.223 – 259

Oviedo C., E.A, Oviedo C., A.I y Vélez S., G.L (2017). *Minería multimedia: hacia la construcción de una metodología y una herramienta de analítica de datos no estructurados. Revista Ingenierías Universidad de Medellín,* vol. 16, No. 31 pp. 125-142 ISSN 1692 - 3324 julio-diciembre de 2017/272 p. Medellín, Colombia

Ramírez, L. A. (2015). *Sostenibilidad o Pensamiento Ambiental*, II Congrés UPC Sostenible 2015 redcolombianafa.org/apc-aa-files/.../2013_05_31_Presentaci_n_OSZ.pdf

Riquelme, J.C., Ruíz, R. y Gilbert, K. (2006). *Minería de Datos: Conceptos y Tendencias*. Inteligencia Artificial: Revista Iberoamericana de Inteligencia Artificial, 10 (29), 11-18

Rojas D. J., Chavarro P. J. y Moreno L. R. (2008). *Técnicas de lógica difusa aplicadas a la minería de datos*, Scientiaettechnica, 3(40). doi: http://dx.doi.org/10.22517/23447214.3095

Stone, A. (2017). *Big Data: all in or all wrong*?, Government Technology, n.º 30, pp. 17-20.

Valencia, H. J. G. (2007). *Conflictos Ambientales: Praxis, Participación, Resistencias Ciudadanas y Pensamiento Ambiental*, Revista Luna Azul, (24), 35-41.

Witten I.H., Eibe, F. and Mark A (2011). *Data Mining, practical Machine Learning Tools and Techniques*, Third Edition, Morgan Kaufman Publishers

Zadeh, L. A. (1965). *Conjuntos borrosos,* Department of Electrical Engineering and Electronics Research Laboratory, University of California, Berkeley, California. Por E. Renedo y S. Guadarrama *Traducido del artículo "Fuzzy Sets", publicado en 1965. relmis.com.ar [23] (1996) "Nacimiento y evolución de la lógica difusa, el softcomputing y la computación con palabras: un punto de vista personal". Psicothema, N° 8, vol. 2, p. 421-429.