Digital Data Visualization with Interactive and Virtual Reality Tools. Review of Current State of the Art and Proposal of a Model

Visualización de Datos con Herramientas Interactivas y de Realidad Virtual. Revisión del Estado Actual y Propuesta de Modelo

Jose Luis Rubio Tamayo
Professor and researcher. Department of Communication Sciences and Sociology (Universidad Rey Juan Carlos - URJC)

Mario Barro Hernández
Postdoctoral Research Fellow (Universidad Nacional Autónoma de México – UNAM) Coordination (Universidad Abierta y Educación a Distancia)

Hernando Gómez Gómez
Associate Professor. Communication Department (Universidad Europea de Madrid - UEM)

Abstract

Massive and open data constitute a burgeoning field of study in the current context. The evolution of technology is, in turn, increasing its degree of interactivity, configuring several scenarios of great complexity in which data is understood on the basis of our interaction with it at different levels. Technologies such as virtual reality or augmented reality present an emerging framework for visualizing, representing and understanding information. Moreover, new disciplines such as interaction design, human-computer interaction, and user experience are needed to optimally configure the representation and design of data interaction dynamics, so that they can be implemented in contexts such as education. This paper reviews the current state of interactive and immersive technology (including virtual reality and alternative reality games) and of open and massive data, to highlight potential projections and propose models of data representation based on factors such as storytelling or user experience. This paper shows the need to develop models for data use and representation in fields such as education and citizen empowerment.

Key Words: *Open access to data - Educational applications - Data visualization - Data science - Human-computer interaction - Interaction design*

Resumen

Los datos masivos y en abierto son campos de estudio con una proyección relevante en el actual contexto. La evolución de la tecnología está, a su vez, incrementando su grado de interactividad, configurando varios escenarios de gran complejidad, en la que los datos son entendidos a partir de la interacción en diferentes niveles. Tecnologías como la realidad virtual o la realidad aumentada presentan un marco emergente para la visualización, la representación y la comprensión de la información. Por otro lado, nuevas disciplinas como el diseño de interacciones, la interacción humano-computadora o la experiencia de usuario, son necesarias para configurar de manera óptima la representación y el diseño de dinámicas interactivas con datos, de manera que sean implementados en contextos tales como la educación. El presente artículo realiza una revisión del estado de las tecnologías interactivas e inmersivas (como la realidad virtual o los juegos de realidad alternativa) y

el estado de los datos masivos y/o en abierto, de manera que se puedan configurar proyecciones y proponer modelos de representación de datos a partir de factores como la narrativa o la experiencia de usuario. El artículo muestra la necesidad de desarrollar modelos en el uso y representación de datos aplicable a campos como la educación y el empoderamiento ciudadano.

Palabras clave: *Acceso abierto a los datos - Aplicaciones educativas - Visualización de datos - Ciencia de datos - Interacción humano-computadora - Diseño de interacciones*

1. Introduction

Today's information society has been configured as a complex scenario in which data usage and information proliferate. Institutions and agencies, on the one hand, have an increasing tendency to leave data open. At the same time, disciplines related to data access and management arise, such as data science, data mining, big data, or deep learning.

In this context, the management and dissemination of data are fundamental elements in the generation of knowledge and in the current scientific context. In this regard, the enormous amount of data that information systems manage requires the development of theoretical models that help us understand the need to optimize data collection, as well as the creation of tools for managing and representing that data. Information and communication technologies, together with increasing levels of interaction through the complexity of technologies, would allow us to develop models not only for representing information, but also for this expanded part of our interface, through immersion and new devices that will allow users to interact and play with it.

On the one hand, we have the open data provided by institutions, much of which needs to be processed in some way in order to find models of interpretation and representation. Since the data is open (a philosophy that applies to non-documentary material, such as mathematical formulas or geographical or demographic information, often consisting of raw data), it is necessary to keep in mind the different technological options that allow us to process it. In this context we can find also terms such as

open access, which refers to distribution and divulgation of data or educational and research resources, in order to improve science thanks to collaboration between peers.

On the other hand, we have massive data management and processing programs for such representation purposes. Another important aspect is the means of representing the data itself. These depend on technological conditions and cultural imaginaries, whereby the same data can be represented in multiple ways and through completely different processes of interaction.

One of the questions arising from this storytelling process concerns the way in which we want to represent the data involved, and which mechanisms we should use to make them interactive. In other words, regardless of whether the data constitutes information of a purely subjective nature, the way in which we want to count it depends on the story we want to build around it, which in turn will depend on associated factors such as user experience or interaction design.

In this respect, we find the factor of the experience we want the user to be immersed in within a virtual world of data. That is, how we imagine the potential scenarios in which the user is immersed in a world of data he understands immediately. How this user experience is integrated, or how the interactions are designed in a context applied, until now, to elements that have been associated with the use of this technology, gives rise to different issues, such as: What representation mechanisms does technology currently offer to represent data and make it interact with the user? And hence, what interaction mechanisms does current technology produce so that the user acquires greater power and more fluid or immersive access to information?

The present paper raises the questions that surround this type of debate, analyzing the elements that could constitute the process of generating and representing data.

2. Purpose

The configuration of a model that, using open data, serves to propose improvements in educational processes is a task that should be considered a result of the

current context, where all those interested in education have more tools than ever to collect or access data. In this respect, the systematization of the process is an eminently desirable goal. Therefore, the following paper is a review of the state of the art regarding the use of interactive and immersive technology as a support for the visualization of data in educational contexts. Specifically, the paper covers:

- Bases for generating models that, based on available information and communication technologies, allow us to optimize the representation of information.

- Future scenarios regarding data visualization, representation, and interaction.

- Configuration of data science applications for designing new educational scenarios.

- Conceptual design, based on the literature, of a series of scenarios that highlight the interrelation between open data, data mining, information management, data visualization and representation, and interactive environment design.

3. Methodology

This research uses different methodological approaches in order to better understand and configure models which help data science to optimize and represent information.

First, we systematically review the literature in order to highlight the aforementioned disciplines and interconnect them in a new framework that encompasses previous approaches and lines of research in disciplines such as data science, open data experiences, big data, deep learning, data visualization, interaction design, human-computer interaction, and immersive environments.

The systematic literature review is normally used to compile and critically analyze some of the important research carried out in a particular area of study. Manuals on systematic literature reviews, such as that of Keele (2007), highlight the importance of synthesizing existing work. They also define certain key features, including the

detection of a significant number of publications in a specific field, which is an indispensable requirement for a quantitative meta-analysis (in Keele, 2007).

Therefore, studies that use the systematic literature review in very specific technological areas, such as interactive and immersive games (Connolly et al., 2012) or software engineering, can implement methodologies in the use and mining of massive data and open data. Furthermore, the various aforementioned disciplines employ a multidimensional approach, in turn integrating studies that highlight the relationship between technology and the user, including interaction design (ID), human-computer interaction (HCI), user experience (UX), and affective computing.

We also propose a structural model based on this systematic review, which addresses the need to take into account the user's relationship with technology in the aforementioned areas, as well as the approaches to data collection and collection modes, so that contributions can be made for the taxonomic categorization of the components. Therefore, the proposed model takes into account the interrelationship between data mining and the open data philosophy of institutions as analogous parts of the same category. The same happens with components such as user experience, interfaces and their potential implications in phenomena such as education or citizen empowerment.

On the other hand, there are studies that define aspects such as hypertextuality, which authors such as Arias-Robles & García-Avilés (2016) apply to areas such as the nature of digital media, thereby expanding the definition of components such as digital interface, user, multimedia content, and structure.

3.1. Literature Review

The systematic literature review for this research started in January 2017. First of all, data and information about the topic was kept in mind. We used keywords such as those aforementioned (big data, open data, open access, data visualization, interactive data, etc.) for a primary search of information. We also defined the questions about the topic and the research field. In this case, questions were focused in the role of data in fields such as research or education. Most of questions were focused in the idea of how to optimize data mining and data representation and visualization.

Different search engines and academic databases where used for this purpose, such as WOS, Google Scholar or SciELO. More than 300 papers about research in data from different approaches where found in a first step. Finally, 60 of the more relevant references about the topic under different approaches where included in this paper. The aim is to offer a broad and in-depth approach to the current context, since this type of review aims to combine as much information as possible with the current evidence.

Subsequently, the criteria were compared and conclusions derived from the knowledge already created by others were given, first, individually in each project or article. Once all the data were obtained, the global conclusion of the revised material was preceded.

The systematic review applied shows the need of introduce various methodologies for use and gather data. In addition, other parts of the process of systematic literature review have been brought together with other domains, such as user-technology relationship, data collection, user experience, interfaces and their future application in education systems. This fact allowed us to assess the quality of the data and described the more important features showing us the current context to develop the proposed model.

After the systematic review of the literature, an alternative to the traditional phases of data management (compilation, processing and representation together with interactive technology) has been done.

The use of open data by public institutions and organizations is the point of intersection between the transparency of public bodies and citizen collaboration (Peled, 2011). In fact, when analogous concepts fall into the same category (within the "data" set), such as big data, data mining, open science, or learning analytics, it actually encompasses a broader phenomenon that has to do with the potential of technologies to explore, organize, and represent data.

But what is meant by "data"? According to authors such as Rosenberg (2013) or Kitchin (2014), in rhetorical terms, data comes before argument or interpretation.

Based on this study, the model is considered as a process that includes related elements and that needs to be observed as a whole in which the intention is to optimize the end user interaction, in order to have an impact on areas or disciplines such as citizen empowerment or education.

In a piece of research that explores data infrastructures, Kitchin (2014) analyzes how the data context is undergoing a process of rapid evolution, and makes a classification that distinguishes open data from big data. In fact, if data mining processes are diametrically different in both open data and big data, the process of synthesizing and further analysis should depend more on the target audience than on the way in which the data is mined.

The main issue is, on the one hand, how data is mined and processed, and, on the other, how it is represented. Both are interdependent and, in turn, dependent on the technological capacity available in any given context. They are, in turn, different processes, especially in relation to representation, since it raises numerous possibilities that highlight the relationship between user, space, technology, information, and image.

Thus, data visualization is an essential feature in the data communication process. Authors such as Marr (2016) emphasize this importance of visualization, stating that it is a key part of the *crucial last step* in data projects. In the context of the enormous amount of data available today, which needs to be managed, this author proposes mechanisms through which this amount of data can be stored and managed in an optimal way. This involves huge analytic operations which aim to mine very specific information from generated and stored digital data.

This same importance of visualization for the democratization of information comes from directly related areas of knowledge such as data design or journalism. Hence, data reporters such as McCandless (2012, 2014) stress the importance of interacting with complex information presented in a simple way, and which is, in turn, easy to interpret. This author also highlights another important factor, which is the creative search for connections and patterns for organizing the information that is going to be presented to the reader (McCandless, 2014). The representation

of information is, therefore, an exercise that calls for a certain degree of abstraction, which in turn requires looking for interconnections between the different patterns that form part of the information. This is highly attractive to the user.

Open access is also relevant in this context. It's defined as the philosophy of distributing and divulgating educational and research resources in on line digital media and without barriers. Many of the grounds in open access to science were stablished by Murray-Rust et al. (2010) in the Panton Principles, a tribute to Panton Arms Pub in Cambridge, often frequented by scientists from the university. Those principles (In Murray-Rust et al., 2010) are four and set some guidelines for designing and distributing educational and scientific resources, such as the use of "recognised waiver or licence appropriate for data" or "publish science in public domain" or stablishing protocols for implementing open access data.

As far as the final data representation process is concerned, this involves several aspects which lead to disciplines such as data visualization (from the point of view of representation), data journalism (from the point of view of dissemination), or attempts to develop interactive or immersive environments. Among these are emerging virtual reality technologies or alternative reality games (ARGs), which would ultimately produce a process of interaction with the user that encompasses more cognitive dimensions, in addition to visual ones.

In this way, the user-information interaction process acquires more cognitive dimensions than purely visual and integrating disciplines such as interaction design, user experience, or human-computer interaction (HCI), as well as affective computation. The point is, in this case, simply to propose models that analyze the interaction of data and information with the user in the way they are represented, so that the data is not only visible but also manipulable, empowering the user through information, and applying it to crucial research areas such as education. Among the literature that deals with these factors we can find, in the early stages, studies by Picard (1995) or, more recently, those of Cambria (2016), on the subject of affective computation. In the case of user experience (UX), we can find, in the present decade, studies of great importance and based on different approaches to the phenomenon, among them those of Garrett (2010), Albert and Tullis (2013),

and Nogueira et al. (2013). Authors such as Büttner et al. (2011) also go in depth in research for data management, defining data in research and focusing also in educational contexts.

More recent studies (Donalek et al., 2014) address the need to find ways of representing data in immersive and collaborative environments. In other words, virtual reality, as a technology, is a medium with particular traits, such as its inter-active and immersive character, which allows us to propose collaborative dynamics that articulate new discourses on the use of data. In any case, there are currently numerous publications and studies related to this convergence between data and immersive environments, namely virtual reality. This is the case of Huang et al. (2001), who combine a range of elements for spatial data analysis, or Helbig et al. (2014), who use 3D visualizations of analytical data to apply them to scientific fields such as meteorology. Wexelblat (2014) also includes it in his work, which reflects on the potential applications of virtual reality.

This review of the literature allows us to consider what immersive and interactive technologies might be like in their present state; a process of transformation in the way we communicate, interact with our peers, and interpret reality. And the quantitative and qualitative leap on the scale of interaction, as Marr (2016) states, that occurs when the user is immersed in a digitally created space with a 360-degree field of vision and simulated movement in three dimensions, should greatly increase the bandwidth of data available to our brains.

3.2. Proposed Model

The literature review presented in the previous section gives rise to the need to formulate and structure models that allow us to visualize and classify, taxonomically, the different phases of data management. These phases would encompass sub-processes such as compilation, processing and representation, which would coexist in a context in which interactive technology evolves at the same time. Chi (2000) presents a taxonomic classification of visualization techniques that shows, in a clear and structured way, data interactions in order to configure new information visualizations. The Data State Reference Model was developed by Chi and Riedl

in the late 1990s (in Chi & Riedl, 1999, and Chi, 2000), with the aim of helping researchers to visualize and deepen understanding of spatial design, as well as to set up a framework in which the people who carry out the implementation can apply a more effective technique of visualization of the information.
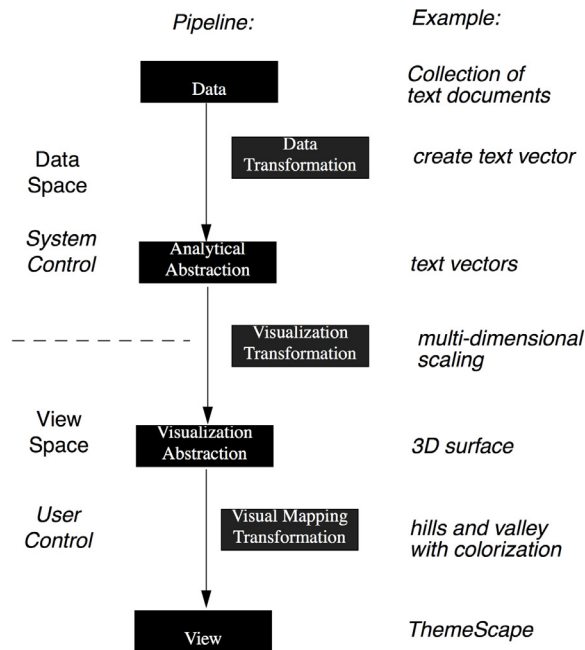

Figure 1: *The Information Visualization Pipeline, taken from Chi and Riedl (1998). For further information on the proposed model, see also Chuah and Roth's Basic Visualization Interaction taxonomy (in Chi and Riedl, 1998).*

Although the model dates from the late 1990s, it is interesting to observe how its configuration serves to develop subsequent tools and methods for interactive data visualization, such as data visualization software (Heer et al., 2005) or studies that explore the visual dimension of data mining and data processing (De Oliveira & Levkowitz, 2003). Recently we find applications of data visualization models in areas as diverse as cognitive science (Goldstone et al., 2014) and development of cognitive frameworks for visualization of information (Patterson et al., 2014), as well as work focusing on the design of complex visualization systems using current technology (Telea, 2014).
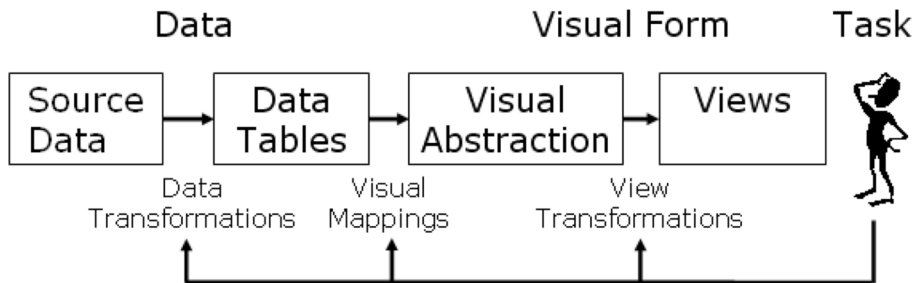
Figure 2: *This schematic diagram model represents the information visualization process. In this case, the model describes the process that goes from the data source, the data tables, the visual abstraction and the representation, in a feedback process that implies the improvement of these process phases. The model has been developed by Prefuse: Information Visualization Toolkit, and can be found on its website as well as in the publications of Heer (2005), who together with Satyanarayan (in Satyanarayan and Heer, 2014), would later design Lyra: An Interactive Visualization Environment.*

Following these lines of research we have, as reference points, studies that consider interaction with data based on the complex framework in which current technology finds itself. Although virtual reality is not a definitive technology, but rather a transition technology (Rubio-Tamayo & Gertrudix Barrio, 2016), the context becomes more complex as technology acquires greater immersive capabilities and increases levels of interaction. Data representation acquires more dimensions from a starting point of static representation as a fixed image, followed by transformation into sounds, manipulation of three-dimensional forms, and even sensations.

In fact, virtual reality, not as a technology, but as a narrative concept involving immersion in a digital world, is currently being studied in order to address this new data representation paradigm. Marr (2016) considers the potential of virtual reality as a technology not only for transforming data visualization, but also for revolutionizing the entire framework. A number of very interesting projects are currently working along these lines, using virtual reality and immersive technologies to propose new models and dynamics of visualization and interaction with data and information (the data set, which ultimately leads to an empowerment of users).

In conceptual terms, our model is based on earlier theories of knowledge management, e.g. the DIKW pyramid, as a way to illustrate the different steps of data and knowledge, which can lead us to build new models. However, we also include critical perspectives and approaches to the model (Frické, 2009), only for the purpose of proposing new updated models.

Based on this combination, the model that arises from the present study entails the following phases:

- Data Collection / Data Mining

- Selective Data Analysis / Processing

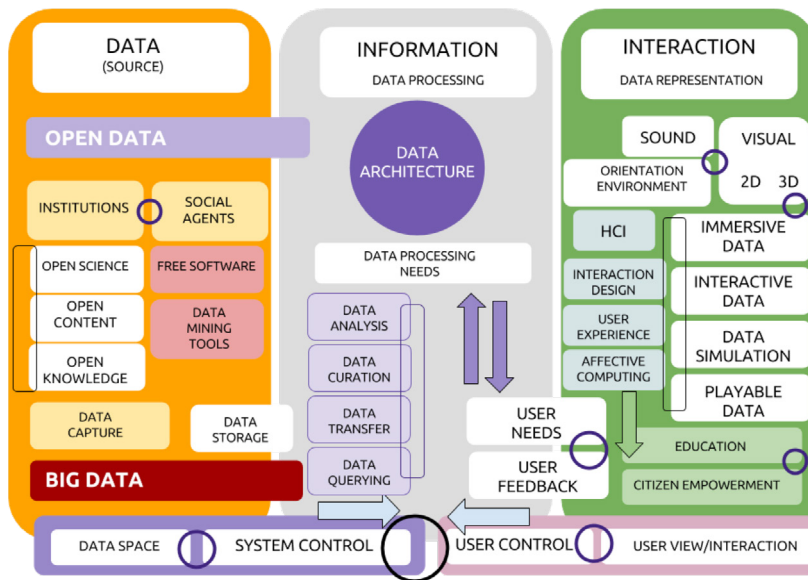Representation / Narratives / Design of Interaction Dynamics

Figure 3: *Model proposed in current research based on review of existing literature and which represents and/or classifies the different components of data gathering and mining, its processing, and design of interaction dynamics. These phases also integrate different disciplines and areas of research in which components related to data management are classified in a taxonomic structure.*

In the case of scenarios involving the use of new models that deepen this relationship between data, information, interaction and immersion, it is worth taking into account factors related to the relationship between user, technology, information or space, such as: the design of interactions and interfaces, user experience, human-computer interaction (HCI) or even affective computing, all related to the potential of user empowerment in the use of technology and applications in fields such as education.

Among the aspects included in models that use immersive and interactive ICT are not only those that refer exclusively to the final interaction of the data with the user, but also those related to the data collection and processing methods. Therefore, data and information would not only be a final representation that could improve the interaction process with the end user, but also process mediated by big data technologies, data mining, open data, or open science, among others, not forgetting the use of data processing tools.

The interest lies in the applications of this context of data use and models based on the one proposed in this exercise, in order to favor learning dynamics. Thus, they could have an impact on the educational context, although it is necessary to carry out studies that address the subject from the perspective of the proposed model and the use of space and technology, through the integration of factors such as interaction design, user experience (UX), and human-computer interaction (HCI).

According to the models presented, data is the object of analysis in a complex process involving the abovementioned phases, which would give rise to new disciplines related to interactive representation. The process would be structured as follows:

- First, the mining/search processes. At this stage, the data is raw and difficult to interpret. There is also the open data offered by institutions, which in this case would facilitate the mining procedure and be easy to access.

- Second, the structuring /processing phase. This phase would involve sorting the data logically so that it can then be structured, as well as looking for patterns and interconnections in order to optimize data representation.

- Third, the representation/visualization process, which seeks to optimize the transmission of this information to groups that are not familiar with the subject matter. Therefore, the aim is to facilitate the process of reading the information. Until relatively recently, and prior to the study of ICT potential, this phase was conclusive.

- The fourth phase corresponds to the development of interaction processes: in this case, different aspects related to body kinetics, narrative, play, and immersion enter into the equation. This is an emerging field of research that involves the abovementioned areas, such as affective computing, human-computer interaction, affective factors, and the potential of digital technology to offer varying degrees of dialogue between the virtual and the physical. This phase acquires greater complexity as technology evolves, and as immersive and interactive potential increases, whereby it is necessary to develop new conceptual models. The approach of a framework that begins with, on the one hand, data mining, and, on the other, open data, needs to generate models which configure an optimization of the user-information relationship at different levels.

Another relevant factor for the development of models is the concept of deep learning, which consists of a set of algorithms that stem from analysis of the image to represent it according to certain perspectives. Studies that address the phenomenon of deep learning (LeCun et al., 2016) claim that it allows machines to learn data representations with multiple levels of abstraction, which are no doubt configured as a contribution to categorize the data flow in different processing layers of processing that allow it to be managed. This phenomenon adds components such as speech recognition or visual patterns through algorithms that evolve together with the technology. Other authors, such as Schmidhuber (2015), apply components such as artificial neural networks in order to explore related phenomena that are not so evident in the use and management of data, such as machine learning or pattern recognition.

Finally, it is important in the study to propose models of taxonomic proposals used in other areas, as has already been applied to interactive and immersive technologies and environments (Rubio-Tamayo, Gértrudix Barrio, 2016). Reference to

this type of study would allow us to visualize more precisely the user's position and relationship with the technology, as well as the immersion capacity and the development of new narratives associated with the phenomenon that we want to communicate with the technology (in this case, the data). In this regard, the importance of developing new models underlines the need to integrate narratives with the data for the configuration of the dynamic, so that it could be applied to some of the related areas, such as data science, in a process of feedback between data mining and the feedback of interactions with the user, as well as a reflection on the fields where it could be applied.

4. Discussion

4.1. Usability Issues

It is common to find usability studies in the context of e-learning, mainly over the last decade, which reflects the validity of the topic in the academic debate about the use of ICTs in education (Wong et al., 2003, Kukulska-Hulme, & Shield, 2004, Granić, 2008, Katsanos et al., 2012). This is probably due to the need to focus on users (their needs and expectations) rather than on the technology and its characteristics. Therefore, it is necessary to address the issue from different perspectives, including information theory, content distribution in ICT mediated contexts, individual and social practices of Internet users, and usability studies. For a definition, we can resort to several authors who have theorized about the concept of usability. Hassan Montero (2002, p. 9) defines it as "the discipline that studies how to design websites so that users can interact with them in the easiest, most comfortable and intuitive way possible."

ISO 9241-11: 1998 defines usability as "Degree in which a product can be used by specified users to achieve concrete objectives with effectiveness, efficiency and satisfaction, in a given context of use" (in Hassan Montero, 2002, p.9).

- Effectiveness: The degree of accuracy with which the system completes the tasks and fulfills the objectives for which it was designed.

- Efficiency: refers to the number of steps that the user must make to complete the task. It considers the expended resources with which the users achieve the goals (resources, time, equipment, and training). Efficiency is the effectiveness achieved considering the resources used. It can be measured in terms of mental or physical effort, time, materials, or financial costs.

- Satisfaction: refers to the comfort, acceptability and positive attitude generated towards the system. Satisfaction has two components: comfort and acceptability.

According to Cooper et al. (2007), *usability* can be defined as the effectiveness, efficiency and satisfaction with which users can achieve specific learning (or learning-related) objectives in a particular environment or with a particular learning tool or resource; and they add that *accessibility and usability* are issues that are intrinsically linked: the lower the level of accessibility of a resource to an individual, the less usable it will be. The same report also points out that accessibility and usability have a direct impact on the effectiveness of systems or resources for all students, but especially for students with disabilities.

4.2. Learning Analytics

Learning Analytics is the measurement, compilation, analysis and reporting of data about students and their contexts, in order to understand and optimize learning and the environments in which it occurs (Siemens and Long, 2011). Overall, Learning Analytics attempts to answer questions about teaching and learning based on data collection and analysis. It is a relatively new discipline. The Learning Analytics concept only came into regular use in 2011, with the definition given at the beginning of this paragraph. For a better understanding, one could begin by breaking down one by one the key terms that make up the definition.

This definition begins with "measurement". Analysis of modern learning is made possible by the fact that education takes place in a context increasingly mediated by technological instruments. As in all scientific disciplines, the instruments used in education broaden our experience, allowing us to *see* more than anyone could

experience directly and with a much greater data storage capacity. Therefore, technological instruments as supports for measuring, controlling and monitoring learning processes are at the heart of Learning Analytics.

The definition continues with "collection" and "analysis". ICTs are the best tools for collecting reliable data on everything we want to measure at any time. It also helps us analyze reality, extract inferences from data sets that are too large or too complex for one person's reach. Therefore, collection and analysis are also supported by ICTs in various ways.

Then we come to a part of the definition that points to the object of Learning Analytics: the data about students and their contexts. What data? In principle, Learning Analytics data can be anything the instruments can record, measure and analyze. This is much more than what we can see with the naked eye and goes beyond quantifying students with categorical or demographic tags to include other elements that intervene in the learning process, such as texts, images, videos, writing patterns, interaction with peers, interaction with instructors, etc. Therefore, the data need not be just numerical realities, given that interesting projects can be designed using qualitative approaches, always based on the collection and processing of a large amount of data.

Who are the apprentices in the definition given? Students, obviously, both in physical and virtual contexts. And why mention the context? Frequently, attempts to quantify learning assume that students are all the same, that they are somehow interchangeable. However, one of Learning Analytics' most interesting application goals points to the possibilities it offers for personalization of teaching and learning. This necessarily includes recognition of the fact that each student in each environment is unique and that effective education should strive to be responsive. Within the context, it is also important to note that students are not the only stakeholders in Learning Analytics. Stakeholders can be anyone who cares about the analysis of higher education: instructors or other members of the educational community who examine how teaching and learning affects ongoing processes.

The end of the definition is a clear statement of the main goal: the purpose of Learning Analytics is to understand and optimize learning and the environments in which it occurs. In other words, understanding teaching and learning is not enough; once we understand, we must act using this knowledge to improve teaching and learning. How? First, by communicating the results of the Learning Analytics project to stakeholders. In the definition this is clearly specified with the term "reporting". And what is it for? For example, for monitoring, performance analysis, prediction, intervention, mentoring, assessment, feedback, adaptation, personalization, recommendation, self-awareness and reflection by the student, instructor, department, faculty, university, etc. (Mattingly et al., 2012, Chatti et al., 2012). The main reason for gathering and processing information is to inform educational practices with a view to carrying out interventions aimed at improving the educational context. In short, data analysis for education is understood as a source of evidence for decision making and a basis for changes and improvements to the system.

5. Conclusions and projection

When analyzing types of interactive and immersive technologies and their evolution, we necessarily consider the potential scope of application and lines of research into how data is used and interpreted, and how interaction systems are designed. In the educational field there are many potential applications for phenomena such as those included in this publication: data mining, access to open data, information management, or even the development of interactive narratives and storytelling with data. At the same time, this projection regarding use takes into account phenomena such as data visualization through interactive and immersive means such as virtual reality, so that scenarios can also be designed at a theoretical and conceptual level, based on these models, allowing them to be implemented in contexts such as education, and boosting the quality of the relationship between the user, the information, the medium in which it spreads, and the space it interacts with.

In this case, interactive and immersive digital technologies are a transition technology that requires studies of, one the one hand, its potential for representation and, on the other, its capacity to interact with the user and generate scenarios in which these can be represented through images, use of space and distances, sound

sources, and even new ways of transforming data into senses or stories. In this regard, it is necessary to integrate the emerging disciplines that analyze factors associated with the relationship between user and technology: UX, HCI, ID, etc.

5.1. The Role of Virtual Reality, Interaction Design and HCI

New emerging disciplines now play the fundamental role in developing new ways of interacting with information. Interactive ICTs based on experiences open the door to the development of new storytelling for representing data. Immersive technologies, such as virtual reality or immersive environments, are tools that can be applied to data representation. They also now have the potential to apply data science and representation in disciplines and research fields such as education.

Therefore, we have already mentioned the applications of virtual reality and their relation to the potential of data visualization. Previous studies and publications such as those already mentioned in this paper (Das et al., 1993, Bryson, 1996) point to multiple possibilities in a field such as the study, processing and visual representation or interactive design of data. Ribarski et al. (1994) proposed using VR for data representation, and Bryson (1996) described the potential of RV for scientific dissemination. Studies carried out in the current decade that present immersive environments as tools for visualizing information (Wexelblat, 2014) opened the door to lines of research into the use of virtual reality as a means of communicating and transmitting information that could be applied to different areas. Such projections include studies by Raya & Sánchez (2014), who concluded that virtual reality could be used to visualize massive flows processed using Big Data techniques, or by the aforementioned by Huang et al. (2001), whose concept of spatial data is combined with analysis and exploration techniques to be applied in immersive environments.

This projection of the study of systems and implementation of factors such as user experience also envisages innovative practices in the future, which combine elements of games to increase usability in immersive environments, as reflected in the studies of Kosmadoudi et al. (2013), or studies of user experience in three-dimensional virtual learning environments, such as Shin et al. (2013). In any case, the UX, ID, and HC are disciplines that from now on will have special significance in the

context of designing interactive experience models that encompass data representation and information communication with different projections and applications.

5.2. Conclusions and Projection in Educational Processes such as Citizen Empowerment

In terms of the applications that could arise from this convergence of data use and representation, immersive environments, interaction design, user experience, and information flow management, two fundamental areas of activity and lines of research are education, on the one hand, and citizen empowerment, on the other.

Regarding education, the relationship between data and education is now a more limited phenomenon, insofar as data is just one of the multiple components of educational processes that include a critical perspective. Accordingly, the models proposed in this paper seek to reflect and explore the many ways in which the phases of the process can be applied to educational dynamics. Therefore, both mining and processing, or data and information representation, and critical reflection on the use of data, are inextricably linked to dynamics and educational processes.

In the entire process of data adaptation, the concept of usability must be prioritized; the simplicity of the technology for the acquisition of user information. The success of E-learning will contemplate both the technological and pedagogical usability. The use of technological resources can´t be above the transmission of knowledge.

The complexity in the process of synthesizing and abstraction of information is a challenge if the main purpose is offer the user the best knowledge. Therefore, in addition to formulating and structuring models to recognize the data management process, a deep analysis and evaluation after the transmission of knowledge (data) is also a priority. The existing models, sometimes, forget this phase although, the Learning Analytics arises with desire to correct and improve the process *sender and receiver*.

With regard to citizen empowerment, the most obvious effect is that the use of open data encourages citizens to directly access information. However, the model design seeks to find new gaps and challenges to foster this citizen empowerment

thanks to the accessibility of information through its becoming interactive. Another direct consequence of this increase in citizen access to data is the promotion of collaborative management processes, as well as an increase in feedback regarding the representation and design of systems that allow us to interact with information. In this way, we propose to reflect on how, in different ways, citizens have the ability to observe and feed back into the different parts of a process, trying to find ways to improve their interaction with information and democratize its access and distribution.

It is necessary to remember that the possibilities offered by the representation of data is enormous. It is a powerful tool where the perceptual process directly affects the cognitive process. The sensory interaction will have a direct consequence in the entire knowledge transmission process.

Data processing increases the ability to solve problems and the concepts big data, open, science, data mining … not only offer solutions in just one way, but its mere application also raises a debate about the process of extraction, treatment and data collection.

The proposals that look to the future in this type of technology revolve around immersive three-dimensional spaces such as those which are already being proposed: Looker Labs about Data Analytics in Virtual Reality or Visualization of High-Dimensional Space, including the *Project NEO: Virtual Reality Data Visualization for Machine Learning* (Bertrand, 2017). That is to say, all this uniting data visualization, big data, machine learning and usability so that the immersive environment is as easy and interpretable as possible.

References

8 Immersive Virtual Reality Visualizations. Virtual Reality Pop. Access on 2017-01-10 [https://virtualrealitypop.com/8-immersive-virtual-reality-data-visualizations-25db54de0c9#.wtffut64m]

Albert, W., & Tullis, T. (2013). *Measuring the user experience: collecting*, analyzing, and presenting usability metrics. Newnes.

Alyssum. Virtual Reality Interactive Data Visualization. Retrieved in 2016-12-27 from [http://www.alyssum.io/]

Arias-Robles, F. & García-Avilés, J.A. (2016): Definiendo la hipertextualidad. Análisis cuantitativo y cualitativo de la evolución del concepto, *Icono 14*, volumen 14 (2), pp. 48-68. [doi: 10.7195/ri14.v24i2.995]

Auer, S., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R., & Ives, Z. (2007). Dbpedia: A nucleus for a web of open data. In *The semantic web* (pp. 722-735). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-540-76298-0_52]

Bertrand, F. (2017). Project NEO: Virtual Reality Data Visualization for Machine Learning. Youtube. Retrieved in 2018-04-12 from [https://www.youtube.com/watch?v=myI4P9C34A0]

Bowman, D. A., & McMahan, R. P. (2007). Virtual reality: how much immersion is enough?, *Computer*, 40(7), 36-43. [DOI: 10.1109/MC.2007.257]

Büttner, S., Hobohm, H. C., & Müller, L. (2011). Research data management.

Cambria, E. (2016). Affective computing and sentiment analysis. *IEEE Intelligent Systems*, 31(2), 102-107. [DOI: 10.1109/MIS.2016.31]

Cooper, M., Colwell, C., & Jelfs, A. (2007). Embedding accessibility and usability: considerations for e-learning research and development projects.

Chatti, M. A., Dyckhoff, A. L., Schroeder, U., & Thüs, H. (2012). A reference model for learning analytics. *International Journal of Technology Enhanced Learning*, 4(5-6), 318-331.

Chi, E. H. H., & Riedl, J. T. (1998, October). An operator interaction framework for visualization systems. In Information Visualization, 1998. Proceedings. *IEEE Symposium* (pp. 63-70). IEEE.

Chi, E. H. H. (2000). A taxonomy of visualization techniques using the data state reference model. In Information Visualization, 2000. InfoVis 2000. *IEEE Symposium* (pp. 69-75). IEEE. [DOI: 10.1109/INFVIS.2000.885092]

Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education*, 59(2), 661-686. [https://doi.org/10.1016/j.compedu.2012.03.004]

De Oliveira, M. F., & Levkowitz, H. (2003). From visual data exploration to visual data mining: a survey. *IEEE Transactions on Visualization and Computer Graphics*, 9(3), 378-394. [DOI: 10.1109/TVCG.2003.1207445]

Djorgovski, S. G. (2012). Novel Approaches to Data Visualization Opening Remarks. *MSR eScience 2012 Workshop*, Chicago, Oct. 2012 [http://research.microsoft.com/en-us/um/redmond/events/escience2012/djorgovski.pdf] [http://dataverse.org/]

Donalek, C., Djorgovski, S. G., Cioc, A., Wang, A., Zhang, J., Lawler, E., ... & Davidoff, S. (2014, October). Immersive and collaborative data visualization using virtual reality platforms. In Big Data (Big Data), *2014 IEEE International Conference* (pp. 609-614). IEEE. [https://arxiv.org/ftp/arxiv/papers/1410/1410.7670.pdf]

Frické, M. (2009). The knowledge pyramid: a critique of the DIKW hierarchy. *Journal of information science*, 35(2), 131-142.

Garrett, J. J. (2010). E*lements of user experience, the: user-centered design for the web and beyond*. Pearson Education.

Goldstone, R. L., Pestilli, F., & Börner, K. (2015). Self-portraits of the brain: cognitive science, data visualization, and communicating brain structure and function. *Trends in cognitive sciences*, 19(8), 462-474. [https://doi.org/10.1016/j.tics.2015.05.012]

Granić, A. (2008). Experience with usability evaluation of e-learning systems. *Universal Access in the Information Society*, 7(4), 209.

Greller, W., & Drachsler, H. (2012). Translating Learning into Numbers: A Generic Framework for Learning Analytics. *Educational technology & society*, 15(3), 42–57.

Hall, D. L., Hall, C. M., & McMullen, S. A. (2017). 20 Perspectives on the Human Side of Data Fusion: Prospects for Improved Effectiveness Using Advanced Human–Computer Interfaces. In *Handbook of multisensor data fusion: theory and practice* (pp. 537-548). CRC Press.

Hassan Montero, Y. (2002). Introducción a la Usabilidad. *No sólo usabilidad*, (1).

Heer, J., Card, S. K., & Landay, J. A. (2005, April). Prefuse: a toolkit for interactive information visualization. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 421-430). ACM. Retrieved on 2017-02-25 from [http://vis.stanford.edu/files/2005-prefuse-CHI.pdf]

Helbig, C., Bauer, H. S., Rink, K., Wulfmeyer, V., Frank, M., & Kolditz, O. (2014). Concept and workflow for 3D visualization of atmospheric data in a virtual reality environment for analytical approaches. E*nvironmental Earth Sciences*, 72(10), 3767-3780. [https://doi.org/10.1007/s12665-014-3136-6]

Huang, B., Jiang, B., & Li, H. (2001). An integration of GIS, virtual reality and the Internet for visualization, analysis and exploration of spatial data. *International Journal of Geographical Information Science*, 15(5), 439-456. [https://doi.org/10.1080/13658810110046574]

iViz. Access on 2016-12-20 [http://www.forbes.com/sites/bernardmarr/2016/05/04/how-vr-will-revolutionize-big-data-visualizations/#463379234ac5] [http://anwell.me/articles/iviz/]

Jennex, M. E. (2009, January). Re-visiting the knowledge pyramid. In System Sciences, 2009. HICSS'09. 42nd Hawaii International Conference on (pp. 1-7). IEEE. [DOI: 10.1109/HICSS.2009.361]

Katsanos, C., Tselios, N., & Xenos, M. (2012, October). Perceived usability evaluation of learning management systems: a first step towards standardization of the System Usability Scale in Greek. In *Informatics (PCI)*, 2012 16th Panhellenic Conference on (pp. 302-307). IEEE. [DOI: 10.1109/PCi.2012.38]

Keele, S. (2007). Guidelines for performing systematic literature reviews in software engineering. In *Technical report*, Ver. 2.3 EBSE Technical Report. EBSE. sn. [doi:10.1145/1134285.1134500]

Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering–a systematic literature review. *Information and software technology*, 51(1), 7-15. [https://doi.org/10.1016/j.infsof.2008.09.009]

Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures and their consequences*. Sage.

Kosmadoudi, Z., Lim, T., Ritchie, J., Louchart, S., Liu, Y., & Sung, R. (2013). Engineering design using game-enhanced CAD: The potential to augment the user experience with game elements. *Computer-Aided Design*, 45(3), 777-795. [https://doi.org/10.1016/j.cad.2012.08.001]

Kukulska-Hulme, A. & Shield, L. (2004). Usability and pedagogical design: Are language learning websites special? In *ED-MEDIA 2004* (pp. 4235-4242).

Laha, B. & Bowman, D. A. (2012). Identifying the benefits of immersion in virtual reality for volume data visualization. In *Immersive visualization revisited workshop of the IEEE VR conference* (pp. 1-2).

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444. [doi:10.1038/nature14539]

Marr, B. (2016). How VR will Revolutionize Big Data Visualization. *Forbes On Line* (May, 4, 2016). Access on 2017-02-12 from [https://www.forbes.com/sites/bernardmarr/2016/05/04/how-vr-will-revolutionize-big-data-visualizations/#c4a2845e1512]

Mattingly, K. D., Rice, M. C., & Berge, Z. L. (2012). Learning analytics as a tool for closing the assessment loop in higher education. *Knowledge Management & E-Learning: An International Journal*, 4(3), 236-247.

McCandless, D. (2012). *Information Is Beautiful*. Collins. New edition (December 6, 2012)

McCandless, D. (2014). *Knowledge Is Beautiful: Impossible Ideas, Invisible Patterns, Hidden Connections*. Harper Design (October 21, 2014)

Murray-Rust, P., Neylon, C., Pollock, R., & Wilbanks, J. (2010). *Panton Principles*, Principles for open data in science. Panton Principles.

Nirvaniq Labs. Retrieved 2017-03-10 from [https://nirvaniq.com/] Prefuse. Retrieved 2017-03-10 from: [http://prefuse.org/doc/manual/introduction/structure/].

Nogueira, P. A., Torres, V., & Rodrigues, R. (2013). Automatic emotional reactions identication: a software tool for offline user experience research. In *Entertainment computing–ICEC 2013* (pp. 164-167). Springer Berlin Heidelberg [https://doi.org/10.1007/978-3-642-41106-9_22]

Patterson, R. E., Blaha, L. M., Grinstein, G. G., Liggett, K. K., Kaveney, D. E., Sheldon, K. C., ... & Moore, J. A. (2014). A human cognition framework for information visualization. *Computers & Graphics*, 42, 42-58. [https://doi.org/10.1016/j.cag.2014.03.002]

Pavlik, J. V. (2015). Transformation: examining the implications of emerging technology for journalism, media and society. *Athens J Mass Media Commun*, 1(1), 9-24.

Peled, A. (2011). When transparency and collaboration collide: The USA open data program. *Journal of the American society for information science and technology*, 62 (11), 2085-2094. [https://doi.org/10.1002/asi.21622]

Picard, R. W. (1995). *Affective computing*.

Raya, L., & Sánchez, A. (2014). Análisis y Visualización de Big Data mediante técnicas de realidad virtual. Big Data Visual Analytics using virtual reality techniques. ArDIn. *Arte, Diseño e Ingeniería* (3).

Rossi, R., & Ahmed, N. (2015, January). The Network Data Repository with Interactive Graph Analytics and Visualization. In *AAAI* (Vol. 15, pp. 4292-4293).

Rubio-Tamayo, J. L. & Gértrudix Barrio, M. (2016): Realidad Virtual (HMD) e Interacción desde la Perspectiva de la Construcción Narrativa y la Comunicación: Propuesta Taxonómica, *Icono 14*, volumen 14 (2), pp. 1-24. doi: 10.7195/ri14.v24i2.965

Ryan, M. L. (2015). *Narrative as Virtual Reality 2: Revisiting Immersion and Interactivity* in Literature and Electronic Media. JHU Press.

Satyanarayan, A., & Heer, J. (2014, June). Lyra: An interactive visualization design environment. In *Computer Graphics Forum* (Vol. 33, No. 3, pp. 351-360). [https://doi.org/10.1111/cgf.12391]

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, 61, 85-117. [https://doi.org/10.1016/j.neunet.2014.09.003]

Schwendimann, B., Rodriguez-Triana, M., Vozniuk, A., Prieto, L., Boroujeni, M., Holzer, A., ... & Dillenbourg, P. (2016). Perceiving learning at a glance: A systematic literature review of learning dashboard research. *IEEE Transactions on Learning Technologies*. DOI: 10.1109/TLT.2016.2599522

Seminario Internacional Escenarios 2020, Analíticas del aprendizaje, hacia una personalización de las ayudas educativas [video completo], Access on december 2016, [https://www.youtube.com/watch?v=LcEwMuze7cM].

Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE review*, 46(5), 30.

Song, W., Wu, D., Wong, R., Fong, S., & Cho, K. (2015, October). A real-time interactive data mining and visualization system using parallel computing. In Digital Information Management (ICDIM), 2015 Tenth International Conference on (pp. 10-13). IEEE. DOI: 10.1109/ICDIM.2015.7381890

Telea, A. C. (2014). *Data visualization: principles and practice*. CRC Press.

Tramullas, J. (2003). Documentos y servicios digitales: de la usabilidad al diseño centrado en el usuario. *El profesional de la información*, 12(2), 107-110.

Underwood, J. (2016). *Immersive Data Visualization with Virtual Reality*. Access on 2016-05-05. [http://www.jenunderwood.com/2016/05/05/immersive-data-visualization-virtual-reality/]

Wexelblat, A. (Ed.). (2014). *Virtual reality: applications and explorations*. Academic Press.

Wong, S. K. B., Nguyen, T. T., Chang, E., & Jayaratna, N. (2003, November). Usability metrics for e-learning. In *OTM Confederated International Conferences" On the Move to Meaningful Internet Systems* (pp. 235-252). Springer Berlin Heidelberg. http://link.springer.com/chapter/10.1007/978-3-540-39962-9_34