Editor's Note

Digital information has redefined the way in which both public and private organizations are faced with the use of data to improve decision making. The importance of Big Data lies in the huge amount of data generated every day, especially following the emergence of online social networks (Facebook, Twitter, Google Plus, etc.) and the exponential growth of devices such as smartphones, smartwatches and other wearables, sensor networks, etc. as well as the possibility of taking into account increasingly updated and more varied information for decision making. [1]

With proper Big Data analysis we can spot trends, get models from historical data for predicting future events or extract patterns from user behaviour, and thus be able to tailor services to the needs of users in a better way.

Using Big Data is becoming widespread among organizations of all kinds. It is a fact that large companies, start-ups, government agencies and non-governmental organizations are gradually being encouraged to use microdata generated by digital devices to operate more efficiently. When these microdata are aggregated they turn into massive amounts of data which require specialized tools and skills to be managed.

The challenge organizations are facing is that the increasing amount of data is too large or too unstructured to be managed and analysed with traditional methods. Think of the data derived from the sequence of clicks from the Web, social media content - tweets, blogs, Facebook wall postings (Facebook alone accounts for more than 1 billion active users generating social interaction content. Google processes on average over 53 thousand search queries per second, making it over 4.6 billion in a single day) - or radio frequency identification systems, which generate up to a thousand times more data than conventional barcode systems (12 million RFID tags – used to capture data and track movement of objects in physical world – had been sold in by 2011. By 2021, it is estimated that that number will have risen to 209 billion. Walmart manages more than 1 million customer transactions per hour). In the World 10.000 payment card transactions are recorded every second. The amount of data transferred over mobile networks increased by 81% to 1.5 Exabyte per month between 2012 and 2014. More than 5 billion people make phone calls, send text messages and surf the Internet with mobile phones. Every day they send 340 million tweets (4.000 per second!). To date they've generated 2.5 trillion bytes of data. However, very little of this information is in the form of rows and columns of traditional databases.

While the increasing amount of data is an undisputable fact, from a social point of view the most relevant issue is the nature and number of problems that the processing of these data is helping to solve. Following Davenport, we could say that although the management of information and data is something that is used in virtually all areas, there are some in which Big Data is used with particular intensity [2]. Among these we include the following: health, politics, finance, security, marketing and business management, the study of social networks, risk analysis, smart cities, human resources, fraud detection, environmental management and education.

Widespread use of Big Data is a result of it having become a problem-solving tool by being able to integrate the traditional tools for managing data with other characteristics of artificial intelligence. Artificial intelligence has been used in several ways to capture and structure Big Data, analysing it to obtain key insights [3].

This special issue is designed with the primary objective of demonstrating what we have just noted: the diversity of fields where big data is used and consequently, how it is gaining increasing importance as a tool for analysis and research. In this sense there are works related to the following topics: digital marketing, optimization of message exchange, sentiment analysis, text analytics, e-learning, financial risk control, forecasting behaviour in video games, energy policy and health.

The first two papers of this special issue are related to Big Data and marketing. Juan Carlos Gonzálvez and Francisco Mochón´s paper, *"Operating an Advertising Programmatic Buying Platform: A Case Study"*, analyses how new technological developments and the possibilities generated by the internet are shaping the online advertising market [4]. More specifically it focuses on a programmatic advertising case study. The origin of the problem is how publishers resort to automated buying and selling when trying to shift unsold inventory. The platform executes, evaluates, manages and optimises display advertising campaigns, all in real-time. The results of this case study show that the platform and discard algorithms incorporated therein are an ideal tool to determine the performance and efficiency of different segments used to promote products. Thanks to Big Data tools and artificial intelligence the platform performs automatically, providing information in a user-friendly, simple manner.

The next paper is also related with marketing and the management of large volumes of data. *"PInCom project: SaaS Big Data platform for ubiquitous marketing using heterogeneous protocols and communication channels"*, written by Juan Manuel Lombardo, Miguel Angel López, Felipe Mirón, and Susana Velasco integrates aspects of cloud computing with the treatment of large volumes of data and the potential and versatility of messaging through various formats and platforms, especially on mobile devices [5]. The importance of this issue arises with the high number of users who have access to these technologies and the information which can be consumed through existing communications networks. The main objective is to offer a cloud service for ubiquitous marketing and loyalty by sending messages using different protocols and communication channels. The platform used is able to handle a lot of traffic and send messages using intelligent routing through standardized protocols while managing information security when it is connected.

"Fine Grain Sentiment Analysis with Semantics in Tweets" presented by the group of researchers of the University of Malaga, consisting of Cristóbal Barba González, José García-Nieto, Ismael Navas-Delgado and José F. Aldana-Montes, is devoted to the study of sentiment analysis [6]. The opinions of Twitter users can be assessed by classifying the sentiment of the tweets as positive or negative [7]. However, tweets can be partially positive and negative at the same time by containing references to different entities within a single tweet. As a result general approaches usually classify these tweets as "neutral". In this paper the authors propose a semantic analysis of tweets using Natural Language Processing to classify the sentiment with regards to the entities mentioned in each tweet. We offer a combination of Big Data tools (under the Apache Hadoop framework) and sentiment analysis using RDF graphs supporting the study of the tweet's lexicon. The work is empirically validated using a sporting event, the 2014 Phillips 66 Big 12 Men's Basketball Championship. The experimental results show a clear correlation between the predicted sentiments with specific events during the championship.

The article by Vidal Alonso and Olga Arranza, *"Big Data & eLearning: A Binomial to the Future of the Knowledge Society"*, focuses on the field of education [8] [9]. The combination of different learning analytical techniques with new paradigms of processing,

DOI: 10.9781/ijimai.2016.360

such as Big Data, will enable relevant information to be accessed by educational authorities and teachers, who in turn will be able to change and optimise current teaching methods. This paper shows a case study where the teacher obtains information about the most popular tools in new learning environments. From this information new strategies of teaching-learning can be set based on student experience so that, looking for greater student participation, the teacher may propose tasks that employ tools preferred by students instead of others less popular. The proposed activities, which involve the use of collaborative tools, stimulate group work, widely regarded as a positive factor in the teaching-learning collaborative process. The use of these tools is highly satisfactory to students, resulting in a more active participation, increased student motivation and a significant improvement in learning.

"Social Network Analysis and Big Data tools applied to the Systemic Risk supervision" written by Mari-Carmen Mochón adopts a different approach to the papers presented so far. It analyses how you could use the huge amount of information generated by financial transactions to combat systemic risk. The importance of this issue is clear and is in line with G20 concerns: the need to strengthen the supervision and control of risk, especially in over- the-counter financial markets [10]. This case study makes a concrete proposal of an analysis methodology. Using Big Data solutions currently applied to Social Network Analysis (SNA), information such as propagation risk can be identified without the need for expensive and demanding technical architectures. This case study exposes how the relations established between the financial market participants could be analysed in order to identify market behaviour and risk of propagation.

The possibilities of using Big Data to address environment and energy problems are discussed in the article by Diego J. Bodas-Sagi and José M. Labeaga, *"Using GDELT Data to Evaluate the Confidence on the Spanish Government Energy Policy"* [11]. This paper analyses the public opinion regarding the Spanish Government's energy policy, using the Global Database of Events, Language, and Tone (GDELT). The aim of the authors is to build sentiment indicators arising from this source of information and, in a final step, evaluate if positive and negative indexes have any effect on the evolution of key market variables such as prices and demand.

"Real-Time Prediction of Gamer Behaviour Using Variable Order Markov and Big Data Technology: A Case of Study", written by Alejandro Baldominos, Esperanza Albacete, Ignacio Marrero and Yago Saez, presents the results found when predicting the behaviour of gamers in commercial videogames datasets. Identifying gaming profiles can help companies to gain a better understanding on how users interact with their games and adapt their products to customers accordingly [12]. Understanding these profiles and learning from them comprise the first step for conditioning the behaviour of gamers to optimize these metrics and to ultimately increase the overall performance of the game. To do so, the company can take an active role so that users from a certain profile can move to other profiles providing better metrics or higher revenues. This paper uses Variable-Order Markov to build a probabilistic model that is able to use the historic behaviour of gamers and to infer what will be their next actions. Being able to predict with accuracy the next user's actions can be of special interest to learn from the behaviour of gamers, to make them more engaged and to reduce churn rate. In order to support a big volume and velocity of data, the system is built on top of the Hadoop ecosystem, using HBase for real-time processing; and the prediction tool is provided as a service and accessible through a RESTful API. The prediction system is evaluated using a case of study with two commercial videogames, attaining promising results with high prediction accuracies.

The use of Big Data in the field of healthcare is experiencing a remarkable growth in such diverse areas as the fight against cancer or the validation of certain drugs [13]. *"Detection of Adverse Reaction to Drugs in Elderly Patients through Predictive Modelling"*, by Rafael San Miguel Carrasco, leverages predictive modelling to uncover new insights related to adverse reaction to drugs in elderly patients. The results of the research show that rigorous analysis of drug interactions and frequent monitoring of patients' adverse reactions to drugs can lower mortality risk.

"Text Analytics: the convergence of Big Data and Artificial Intelligence" by Antonio Moreno and Teófilo Redondo focuses on the study of what is known as text analytics, i.e. the analysis of the text contained in emails, blogs, tweets, forums and other forms of textual communication [14]. Text Analytics has produced useful applications for everyday use. The authors discuss the following three: Lynguo, IBM's Watson, and IPsoft's Amelia.

Francisco Mochón

Juan Carlos Gonzálvez

References

[1] Aldana J, Baldominos A, García JM, Gonzálvez JC, Mochón F, Navas I; (2016). "Introducción al Big Data", *Editorial Garcia-Maroto editores S.L.*

[2] Thomas H. Davenport. (2014). Big Data at Work: Dispelling the Myths, Uncovering the Opportunities. *Harvard Business Review*.

[3] Daniel E. O'Leary. (2013). "Artificial Intelligence and Big Data", *IEEE Intelligent Systems, vol.28, no. 2, pp. 96-99, March-April 2013, doi:10.1109/MIS.2013.39.*

[4] Thomas H. Davenport and Julia Kirby. (2015). Beyond Automation. *Harvard Business Review. June, 2015.*

[5] Viktor Mayer-Schönberger,Kenneth. (2013). CukieBig Data: A Revolution that Will Transform how We Live, Work, and Think. *Houghton Mifflin Harcourt*.

[6] Bing Liu. (2010). Sentiment Analysis and Subjectivity. In Handbook of Natural Language Processing, Second Edition, (*editors: N. Indurkhya and F. J. Damerau*).

[7] F. Mochón y O. Sanjuán. (2014). A First approach to the implicit measurement of happiness in Latin America through the use of social networks. *International Journal of Interactive Multimedia and Artificial Inteligence. March 2014.*

[8] George Siemens and Phil Long. (2011). "Penetrating the Fog: Analytics in Learning and Education" *EDUCAUSE Review, v46 n5 p30-32, 34, 36, 38, 40 Sep-Oct 2011.*

[9] Anthony G. Picciano. (2012). "The Evolution of Big Data and Learning Analytics in American Higher Education." *Journal of Asynchronous Learning Networks, v16 n3 p9-20 Jun 2012.*

[10] Daning Hu, Gerhard Schwabe and Xiao LiEmail (2015). "Systemic risk management and investment analysis with financial network analytics: research opportunities and challenges." *Financial Innovation. June, 2015 1:2.*

[11] Claudia Vitoloa, Yehia Elkhatibb, Dominik Reusserc, Christopher J.A. Macleodd, and Wouter Buytaerta.(2015). Web technologies for environmental Big Data. Environmental Modelling & Software. *Volume 63, January, pages 185–198.*

[12] Steven Rosenbush and Michael Totty. (2013). "How Big Data Is Changing the Whole Equation for Business." *Journal Reports: Leadership. The Wall Street Journal, March 10, 2013*.

[13] Hsinchun Chen, Roger H. L. Chiang and Veda C. Storey. (2012). Business Intelligence and Annalytics: from big data to bog impact. *MIS quarterly, 2012 http://hmchen.shidler.hawaii.edu/Chen_big_data_MISQ_2012.pdf.*

[14] Vishal Gupta and Gurpreet S. Lehal. (2009). A Survey of Text Mining Techniques and Applications. *Journal of Emerging technologies in web intelligence,vol. 1, Nº. 1 August 2009.*