Line 2

Technology: Galvanizing the University-Company-State Relationship for Regional Development

Abstract

This line is a compendium of texts that examine the UCS partnership's approach to innovation from a technological perspective. The texts point out that accelerated technological advances are outpacing the ability to control them and that we must therefore make sense of the complexity to discover important relationships that facilitate knowledge and its transfer. It is acknowledged that complexity sciences have emerged as a knowledge revolution, presenting theories, models, concepts, methods, and logics in the search for answers that should be addressed by different disciplines. The impact of innovation hubs increases the technological capacity of research groups, establishes learning communities focused on specific topics, and widens the landscape and network of relationships among the region's stakeholders. It also builds the intellectual capital of the organizations involved in the hubs, while technological advance encourages the training of businesspeople.

The texts that make up this section offer new reflections on and questions about the close relationships between communication and information and communications technology (ICT), the wellbeing of society, the relationship between the university and the business sector, competitiveness, and regional growth. The articles highlight and consider ICT as the driving force of processes, making it possible to transform both the apparatus of production and society in general. Under this technological perspective and context that characterize the 21^{st} century, the texts forming this line allow the incursion of new paradigms and approaches into this field for deliberating on and questioning the technological development and regional growth based on sustainable human development. They also consider and propose technological advance as the chief motive for streamlining regional transformation processes.

Keywords: complexity, innovation, research, University-Company-State relationship, technology.

Línea 2

La tecnología: motor que dinamiza la relación Universidad-Empresa-Estado, para el desarrollo regional

Resumen

Esta línea compendia cinco capítulos cuyos textos se presentan desde la perspectiva de la alianza Universidad-Empresa-Estado de cara a la innovación, desde la óptica de la tecnología. En ellos se señala que el acelerado avance tecnológico está superando la capacidad para controlarlo, de ahí la necesidad de dar un sentido desde la complejidad, para encontrar relaciones significantes que permitan el conocimiento y su transferencia. Se reconoce que las ciencias de la complejidad emergen como una revolución del conocimiento, las cuales presentan teorías, modelos, conceptos, métodos y lógicas en la búsqueda de respuestas que deben ser abordadas desde diversas disciplinas. La proyección de los núcleos de innovación amplía las capacidades tecnológicas de los grupos de investigación, lo que permite establecer comunidades de aprendizajes enfocados en temas específicos, para ampliar la red de relacionamiento de actores de la región, fortalecer el capital intelectual de las organizaciones que forman parte de los núcleos; además, el avance tecnológico favorece el proceso de formación de empresarios.

En los textos que integran esta línea se plantean nuevas reflexiones y cuestionamientos acerca de la estrecha relación entre las denominadas TIC y la comunicación, el bienestar de la sociedad, el relacionamiento de la Universidad con el sector empresarial, la competitividad y el crecimiento de las regiones. Se destacan y consideran las TIC como el motor que dinamiza los procesos en una oportunidad de transformación tanto del aparato productivo como de la sociedad en general. Bajo este aspecto y contexto tecnológico que caracteriza al siglo XXI, los escritos presentes en esta línea permiten la incursión en este campo de nuevos paradigmas y enfoques para deliberar y cuestionar sobre el desarrollo tecnológico y el crecimiento de las regiones a partir del desarrollo humano sostenible. Consideran y proponen el avance tecnológico, como el motor que dinamiza los procesos de transformación regional.

Palabras clave: complejidad, innovación, investigación, relación Universidad-Empresa-Estado, tecnología.

Chapter 6 / **Capítulo 6**

Big Data in the Complexity of the University-Company-State System

El *big data* en la complejidad del sistema Universidad-Empresa-Estado

Juan Carlos Alvarado-Pérez

Abstract

Big data is a new concept that has emerged in response to current advances in electronics and informatics. These advances have contributed to the rapid formation of mountains of data on a disproportionate scale, outstripping the ability of conventional tools to capture, manage, and process them in a reasonable time frame. This data is scattered and originates from different sources, which have the potential to be combined, giving rise to inconsistent and unpredictable compositions.

Complex systems, such as the University, the Company, and the State, face various problems, about which there exists a great quantity of data but little information. It is therefore necessary to bring order to the chaos, finding significant relationships to generate new knowledge. For this reason, the tools must evolve in tandem with the growth of data, resulting in novel techniques that harness the potential of computational resources, such as data mining, and integrate these techniques in synergy with others that are adapted to the qualities of human analysis and interpretation, such as visual analytics. In other words, artificial and natural intelligence must be combined to create a mechanism suitable for discovering knowledge and attaining a holistic worldview.

Keywords: big data, complexity, company, state, intelligent systems, university.

Resumen

Big data es un nuevo concepto que emerge en respuesta a los actuales avances electrónicos e informáticos, los que han contribuido a formar montañas de datos generadas a un acelerado ritmo y que han alcanzado dimensiones exorbitantes, superando la capacidad de las herramientas convencionales para ser capturados, gestionados y procesados en un tiempo razonable. Además, dichos datos están dispersos y proceden de diversas fuentes, las cuales tienen la capacidad potencial de combinarse, propiciando composiciones inconsistentes e impredecibles.

Los sistemas complejos, como la Universidad, la Empresa y el Estado, enfrentan diversos problemas, sobre los que existen gran cantidad de datos pero escasa información, lo que hace necesario dar un sentido a todo este caos, encontrando relaciones significantes para generar nuevo conocimiento. Por esta razón, las herramientas deben evolucionar paralelamente al crecimiento de los datos, haciendo emerger novedosas técnicas que utilicen las potencialidades de los recursos computacionales, como la minería de datos, y los integren de manera sinérgica con otras técnicas que se adapten a las cualidades de análisis e interpretación humana como la analítica visual, es decir, la suma de la inteligencia artificial y la inteligencia natural como un equipo idóneo para descubrir conocimiento y concebir una visión holística del mundo.

Palabras clave: *big data*, complejidad, empresa, estado, sistemas inteligentes, universidad.

Pages 97-115

Author profile / Perfil de autor

Juan Carlos Alvarado-Pérez
Currently undertaking a PhD in Computer Science and Automation. Master's degree in intelligent systems. Professor in the School of Medicine at Universidad Cooperativa de Colombia, Pasto campus.
E-mail: juan.alvarado@campusucc.edu.co

How to cite this work / ¿Cómo citar?

APA

Alvarado-Pérez, J.C. (2014). Big Data in the Complexity of the University-Company-State System. In Hernández Arteaga, I. & Pemberthy-Gallo, L.S. (Comps.), *University-Company-State: Towards a Culture of Research and Innovation. Second Conference of Business Innovation in Cauca and Nariño* (pp. 97-115). Bogotá: Editorial Universidad Cooperativa de Colombia - Red UREL.

Chicago

Alvarado-Pérez, Juan Carlos. "Big Data in the Complexity of the University-Company-State System." In *University-Company-State: Towards a Culture of Research and Innovation. Second Conference of Business Innovation in Cauca and Nariño*, comps. Isabel Hernández Arteaga and Luz Stella Pemberthy-Gallo. Bogotá: Editorial Universidad Cooperativa de Colombia - Red UREL, 2014.

MLA

Alvarado-Pérez, Juan Carlos. "Big Data in the Complexity of the University-Company-State System." *University-Company-State: Towards a Culture of Research and Innovation. Second Conference of Business Innovation in Cauca and Nariño*. Hernández Arteaga, Isabel and Pemberthy-Gallo, Luz Stella. (Comps.). Bogotá: Editorial Universidad Cooperativa de Colombia - Red UREL, 2014, pp. 97-115.

Gigantic volumes of data: Big data

The term big data implies, in itself, complexity. It is a new concept that has emerged in response to the great volume of data that has reached unmanageable dimensions (Schmarzo, 2011); that is, the data has exceeded the ability of conventional software to capture, manage, and process it in a reasonable time frame, exposing the limitations of the underlying infrastructure. Big data is therefore any attribute that represents a challenge to the limited abilities of systems or to the needs of the State, companies, universities, health, or any other area capable of generating knowledge.

Current advances in technology, such as web platforms, cloud computing,[1] and social networks, are driving the advance of big data, whose high demand inspires better web designs, establishing a cycle of continual improvement along with constant, dynamic updates. This places big data in a state of constant growth, and as time passes its data evolves and becomes more precise, generating data about data, or metadata. Big data is thus converted not only into a source of knowledge, but also a major business opportunity in itself.

Big data not only refers to the size of the data; it also considers the rapid rate at which data is generated and the speed at which diverse sources of data are created (Schmarzo, 2011). These sources may be independent, and they potentially have the ability to interact among themselves, giving rise to inconsistent and unpredictable compositions. In other words, it is possible to generate one gigantic database from a combination of small databases that have some type of relationship. Not all data sources have an obvious relationship, however, and consequently not all big data is the same from a structural standpoint. Some big data has a well-defined format, such as the relational transactions supplied by a database management system (DBMS), while other big data may be no more than blog entries that contain text, images, voice, and video, kept in the same storage space, like the files generated in social networks. These are an enormous source of data, with 96% of people born between 1983 and 1995 belonging to these social networks. Facebook alone currently has 900 million users (*El País*, 2012), making it the number one activity on the internet.

It took radio 38 years to reach 50 million users. Television took 13 years to do the same, the internet 4, and the iPod 3, whereas Facebook added 100 million users in only 9 months (Zambrano Reyna, 2011). But it is not only the social networks that contribute to this explosion of bytes; in platforms such as YouTube, 500 years of video are viewed every day, and 1 billion videos are uploaded every year. In 60 seconds, 600 new videos are created, 100 new LinkedIn accounts are opened, 13,000 iPhone

---

[1] New IT paradigm that allows computer services to be offered over the internet.

apps are downloaded, 694,445 searches are conducted, 98,000 tweets are posted, 168 million e-mails are sent, and 379 Skype (EMC) voice calls are made.

The rise of these technologies has been dizzying, and they show no signs of slowing down. Owing to mobile platforms, big data is everywhere. For this reason, the combination of unstructured data with structured data can provide abundant information for the company, the university, and the State. Consequently, the opportunities to improve understanding of society through the use of big data are almost limitless, offering many competitive advantages, generating highly positive social and economic changes, and, finally, enabling richer personal and work lives.

The digital universe has been forming for a long time. In 1948 Claude Shannon introduced the term bit, designating it as the measure of a unit of information (Heredia, n.d.). The vast quantity of data collected over time in different forms and in different areas such as health, industry, the company, and the State, has increased in volume due to current advances in electronics and informatics. These include sensors, satellites, magnetic bands, and the internet, among many others, and contribute independently to the creation of enormous mountains of data.

In 2012 alone approximately 2.5 quintillion bytes (2.5×10^{18}) were generated daily. The network is fed more than 15 petabits[2] daily, according to IBM; the total information stored globally reached 2.7 zettabits[3] in 2012, according to IDC, and it doubles every 18 months. In 2020 it is estimated that the accumulated information will be 35 zettabits (Gantz, Reinsel, & Shadows, 2012), which is 50 times greater than the number of grains of sand on all the Earth's beaches, keeping in mind that there are 700.500.000.000.000.000.000 grains of sand in total (*El País*, 2011); nevertheless, it is predicted that the volume of information will be 50 times greater at the end of the decade.

However, the ability to interpret, analyze, summarize, and manage such volumes of data is improving at a much slower pace. Traditional tools such as spreadsheets; statistics; flat, static, and conventional diagrams or graphics; and the conventional analysis of business intelligence are no longer sufficient, as stated by Leonardo González Barceló[4] (Rúa, 2012). It is at this point that the tools must evolve in tandem with the growth of data and in accordance with experts' needs. It is therefore necessary for new techniques to emerge that utilize the potential of

---

2 A petabyte (PB) is a unit of information storage equivalent to 1,0005 bytes = 1.000.000.000.000.000 bytes.

3 A zettabyte (ZB) is a unit of information storage equivalent to 1021 bytes.

4 Currently IBM's Sales Leader of Big Data and Data Warehousing for Latin America.

computational resources and integrate them in synergy with the characteristics of human analysis and interpretation; that is, artificial and natural intelligence must be integrated to form a suitable mechanism for discovering knowledge. What is needed are business analytics.

Accordingly, new roles and even professional profiles are currently being created based on the concept of big data. These include the manager of online reputation, responsible for creating a hierarchy of webpages by applying cybermetrics. Another profile is that of the knowledge manager, who is akin to a data scientist, attempting to transfer experience and existing knowledge among members of an organization so that they can be made available as a resource to others in the organization.

The new knowledge society, in which information is extremely valuable and can be transmitted almost instantaneously, has created a world rich in data but with information that is very scattered. This information can only be made useful through intelligent processes that provide access to effective solutions for solving the problems of the modern world and that aid in the decision-making of our leaders, businesspeople, and scientists. The collection, analysis, and compression of big data implies greater gains and savings for both the company and the State, and it is here that the university can efficiently coordinate tools, processes, and strategies to understand the contexts in which problems arise. This, in turn, allows managers and leaders in business and politics to devise solutions taking into account the nature of the problem at hand, as otherwise they will surely make the situation worse. Accordingly, decision makers need new concepts, approaches, and instruments that allow them to be more proactive and define how the future will look rather than simply reacting to the unexpected consequences of the past.

Complexity

Complexity sciences have emerged as a knowledge revolution, presenting a great number of theories, models, concepts, methods, and logics (Zoya & Aguirre, 2011) to explain why situations become complex, why systems tend towards chaos, and why they generate self-regulating forces. The answers to these questions represent the key pillars of complexity, providing multiple responses that generally do not reach a consensus, and must therefore be looked at from different disciplines. Furthermore, complexity attempts to explain the order and chaos that affect a system, or in other words, how order becomes chaos and vice versa, how subtle changes can generate drastically different results, and how phenomena can make systems evolve and become more complex, acquiring varying degrees of difficulty (Maldonado & Gómez Cruz, 2010).

Adam Smith's notion of the "invisible hand" (Gache & Otero, 2010) is well known as the phenomenon which gives rise to a stable and efficient economy where direct government intervention in natural macroeconomic dynamics has generally had adverse affects; however, the same economy and the agents that interact within it (supply and demand) have the capacity to regulate themselves, as if there were an invisible force. This self-regulating force has been widely studied in complexity theory, which is applied in natural sciences such as physics and statistics. The concepts developed in natural sciences can also be employed in social sciences, as there is something akin to a physics of society in which social and economic systems possess some characteristics that appear similar to the properties of certain physical and biological systems.

For example, social systems tend to develop hierarchical organizations and form groups similar to schools of fish or herds of animals (Ball, 2012); likewise, these systems display behavior similar to neural networks or cellular biology. There may also be parallels with non-living systems such as avalanches of snow or piles of sand, in which each particle exerts an influence on the next, creating a cascade effect, which is very common in complex systems such as social, biological, and technological systems. This is the case for outbreaks of panic in the markets, fault propagation in the electricity grid, and the spread of epidemics of infectious and contagious diseases.

Another example of a physical concept similar to a social system is that of thermodynamics (Zoya & Aguirre, 2011), which studies the exchange of heat energy between one system and another until equilibrium is reached. The energy transmitted from one particle to another is similar to the information that can be transferred from one individual to another in a social system (Ball, 2012), generating situations of speculation until an information context is established that would be the state of equilibrium, the same way it occurs in the stock exchange or economic systems.

Chaos and complexity theory are present in all events that govern the very nature of the laws of the universe (Reynoso, 2007), and thus it has been studied from different perspectives and fields. But the fields that have concerned themselves most with understanding this theory are the natural sciences, such as physics, mathematics, and statistics, in an attempt to make sense of the apparently random nature of the world.

"The invisible forces" that govern systems, such as entropy,[5] negenthropy,[6] and forces of attraction and repulsion, among others, affect the dynamics particular to systems. These dynamics, added to the infinite variables and possible values, make it very difficult to predict the behavior of a complex system, that is, of a system that has different levels of organization and that has evolved to more unstable states, such as social systems, for example. For this reason, the unpredictability of these systems is greater than for natural systems, as the number of variables involved tends towards infinity, meaning that the results of an experiment in apparently similar conditions can differ drastically, as explained in the butterfly effect theory (Stewart, 2007). However, it is possible that many of the concepts applied in natural sciences also apply to social contexts, such as the State, the company, and the university, among other areas, as they display similar properties and characteristics, such as the forming of hierarchies, groupings, self-regulation, etc.

Nature, intelligent and independent, is capable of self-regulation and self-awareness in order to choose the most appropriate action through the interaction of its elements and not by exercising governing forces that alter its natural behavior. This same concept of complexity can be extended to social systems, as Thomas Schelling reflects on in his writings[7] (Salazar, 2007; Schelling, 1981 & 2008). For example, some functions such as collective movements and forms of group organization are highly sensitive to the details of how people interact, and they are determined by the very fact of these interactions. This is the case for traffic light systems, which in their traditional form have a governing model that gives the right of way using a static method for specific time intervals, making the system inefficient.

If complexity theory were used, however, with the elements that interact within the system being those that control it, the situation would improve dramatically. Obtaining information about traffic flow through sensors would allow each traffic light to respond adaptively to traffic flow conditions at all times, optimizing the right of way and giving priority to more congested routes and the intersections with greatest traffic pressure. The traffic itself would thus control the lights, and not vice versa, just as modeled in European and U.S. cities with the TRANSIMS[8] system, which has been used to plan road networks.

---

5 Based on the second law of thermodynamics, which posits that the loss of energy in systems leads them to consume themselves, become disorganized, and die.

6 Opposite process to entropy, necessary to compensate for the system's process of degradation bringing in energy or information from its surrounding environment.

7 Nobel Prize winner in Economics 2005 and author of *The Strategy of Conflict*, considered one of the 100 most influential books since 1945.

8 Transport analysis simulation system developed by NASA under an open source agreement.

Accordingly, it is the elements within the system that should regulate and govern it, breaking the authoritarian paradigm and abandoning top-down hierarchical control to give rise to efficient forms of behavior. Ruling forces are necessary, but they should be focused on other objectives: not on directing the system itself but on ensuring the resources that allow the system to function independently. In the case of traffic lights, for example, these resources would be sensors and mechanisms that facilitate dynamic interaction in the system. There are many benefits, as TRANSIMS shows, with results reflected in reduced traffic and accident and pollution rates, leading to the economic benefits of using less fuel, among other collateral gains.

Social behavior is extremely complex and arises spontaneously through the interaction and influence of individuals in a system. For example, in open green spaces such as public parks, there is a tendency for people to walk where others have trodden down the grass, owing to the psychological impulse to follow paths. When these individuals are independent entities that can make decisions, such as animals, people, institutions, etc., they are often referred to as agents, and they continually change and are changed by a great number of variables that alter their individual behavior, which, in turn, modifies the behavior of the system in general.

Consequently, results usually depend on a series of different contingencies that make them exceedingly sensitive to, and dependent on, random factors, making social systems extremely difficult to predict from an external perspective. In other words, small individual decisions affect general behavior, just as a school of fish or a flock of birds moves in an apparently synchronized way in the presence of a predator, as if someone was directing them, when in fact it is nothing more than the will of each agent, whose movement is transmitted to the nearest neighbor, and the next, and so on, through simple rules with the aim of avoiding collision among individuals. This makes a system that can involve a great number of beings behave like a single being in itself, and it implies that the interaction of many agents does not necessarily lead to chaos and unpredictability.

This ability of complex systems to adopt ordered behavior patterns is often called self-organization[9] (Ball, 2012). This is not imposed from above, that is, rather than agents following a leader, self-organization arises spontaneously from the bottom up. Likewise, the laws in a social system emerge dynamically based on the social, cultural, and economic context in which they are generated.

---

9 Spontaneous process in which order and coordination arise from interaction among the components of an initially disorganized system that is not controlled by any agent or internal or external subsystem.

Human behavior involves emotion, creativity, and intention, making the human social system highly unpredictable; thus, these systems are the most complex that we know of. The key principle presented here is that the concept described above provides a reference for applying complexity in social systems in any area where there is data that can be converted into information and, subsequently, into knowledge of the invisible rules that govern the systems. To do this, modern tools are needed that uncover the information hidden in the gigantic mountains of data.

As a result, new techniques and tools that facilitate this task have recently emerged. One example is KDD[10], a new and innovative field and process that essentially involves combining different techniques of discovery and knowledge such as data mining (López & Herrero, 2006) and visual analytics (Butz et al., 2009), among other intelligent and independent systems. These allow a robust analysis of data and events in order to identify patterns and generate mind maps and, subsequently, new knowledge. The patterns discovered are presented in the form of rules or functions so that the analyst can interpret them with the aim of understanding and even predicting the world we have created.

This process could be applied to chaotic and complex systems to discover the forces that govern them, in the form of rules of engagement, cluster, or classification (García, Sánchez, & García, 2006; Pereira, Millán, & Machuca, 2006), allowing events to be made sense of. These new support technologies help in understanding and predicting behavior of complex systems, as well as—literally—orienting them towards optimal states. In other words, the important thing is not to impose solutions, but rather to adjust the rules that govern interaction among agents so that forms of behavior arise spontaneously and efficiently. This has been the contribution of various universities to companies and the State, developing tools such as VisMineKDD,[11] Weka,[12] Orange,[13] RapidMiner,[14] JHeWork,[15] knime,[16] and many more.

---

10 Knowledge discovery in databases.

11 Developed by Universidad Cooperativa de Colombia and Universidad de Salamanca, Spain.

12 Developed by the University of Waikato, New Zealand. Available at http://www.cs.waikato.ac.nz/~ml/weka/

13 Developed by the University of Llubljana: http://orange.biolab.si/

14 Previously named YALE, it was developed by the University of Dortmund: http://rapid-i.com/content/view/181/190/

15 A free KDD tool designed for scientists, engineers, and students: http://jwork.org/jhepwork/

16 Konstanz Information Miner, developed by the University of Konstanz, Germany, http://www.knime.org/

The tools presented here can be used to analyze data generated not only by the company and the State, but also by multiple areas such as health, culture, and religion. This, in turn, allows the assimilation of the general context of the complex behavior of a society, in which the knowledge generated is of benefit to all. The company can increase their profits by identifying new markets and potential customers, and detecting faults, fraud, and anomalous processes, among other benefits. The State can create greater social wellbeing by understanding the internal workings of the system, applying this knowledge when developing strategies to ensure the correct functioning of the system, without regulating it in an authoritative way. In the area of education, dropouts and poor academic performance could be avoided, and the areas of knowledge most needed in the region could be identified. The needs of our teachers and students could also be understood.

The scope of this model encompasses not only the university, the company, and the State, but also any facet of society in which the data of one affects the data of another. For example, if a business such as a mall is located in the outskirts of a city, it will surely affect mobility around it, and this should be taken into account by the government when planning the road network in the area. Likewise, if government officials change the school calendar, this will affect the companies that provide school supplies as well as the education sector. Therefore, information cannot be understood or interpreted in isolated contexts, but instead must be understood from a holistic perspective where these sets of data interact. For this reason, precise and up-to-the-minute information is required; the problem is that the information is stored in a scattered way and in both structured and unstructured formats, making it unavailable. Nevertheless, progress has been made in this area, and many private companies, state institutions, and even governments in various countries have started creating their own data warehouses,[17] allowing them to monitor information and manage it efficiently. The European Union is a pioneer in this area and has invested heavily in research, generating immeasurable benefits.

Common contexts: University-Company-State

There are a great number of common situations in which the company and the State can and should share information to generate new knowledge and learn more about our social systems. Some of these situations are presented below.

---

17 A summarized repository of data from multiple sources that can be searched along specific lines in simple ways to support the decision-making process.

Situation 1. Georeferenced movement

In recent years, modern techniques of data acquisition such as global positioning systems (GPS), radio-frequency identification (RFID), camera networks, wireless sensors, and cell phones have seen great quantities of geographic data obtained in the form of trajectories (Vieira, Bakalov, & Tsotras, 2009). The popularity of these technologies and the ubiquitous nature of mobile devices suggest that the amount of time-space data will rise sharply in the near future.

But despite the growing demand for instruments to manage geographic and temporal information, there are few tools available for conducting a satisfactory study of time-space data sets. The natural complexity of these data sets, their accuracy, their privacy, and their enormous volume have hindered an efficient analysis of georeferenced data. New and efficient techniques are therefore urgently needed to help extract valuable information from these masses of data.

Recently, new interest has emerged in discovering group patterns or patterns of common behavior. An example of this type of pattern is movement in a herd, convoy, or flock (Romero, 2011), defined as a group of objects in associated movement, that is, with a predefined distance between them and moving in one direction for a certain amount of continuous time, carrying out collective negotiation.

Much like fish and birds, humans tend to move in groups (Romero, 2011), adopting characteristics of common movement. Detecting flock-like patterns is an extremely important field owing to the characteristics of the objects of study (animals, pedestrians, vehicles, or natural phenomena), which help establish the interaction among objects over a relative time (Gudmundsson, 2004). Organized movement can arise from many individual decisions without a coordinating leader, emerging from the system's individuals themselves. These collective movements have become an emblem for complexity science, in the sense that apparently intelligent and coordinated behavior can stem from simple rules of interaction.

Due to the increasing availability of spatial databases, different methodologies have been explored in order to find useful and hidden information in this kind of data, creating a new perspective of how diverse entities move in a spatial context. These methodologies have proved useful in fields as varied as sports (Iwase & Saito, 2002), socioeconomics (Frank, Raper, & Chaylan, 2001), animal migration (Dettki, Ericsson, & Edenius, 2004), and security (Makris & Ellis, 2002).

Pedestrian movement in Beijing is a real-life example of a case study based on a set of data of georeferenced trajectories that was collected during Microsoft Research

Asia's GeoLife[18] project. The data contains information about 165 anonymous individuals and was captured from April 2007 until August 2009. Locations were recorded by different GPS devices or smart phones; the majority of these provided a high sample rate with 95% of readings registered at intervals of between 2 to 5 seconds, or one reading every 5 to 10 meters.

Although the data set included locations across more than 30 cities in China and even in the Americas and Europe, the majority of data was captured in Beijing. The data set contains information on the latitude, longitude, altitude, date, and time of recording for a wide range of the individuals' movements, not only routines such as going to work and returning home, but also recreational and sports activities. It is important to mention that the participants used varying modes of transport during the study, so movements may refer to traveling by car, on foot, or in public transport. The data set has been used in previous research, identifying mobility patterns and preferred places (Zheng, Zhang, Xie, & Ma, 2009). This information is freely available at the GeoLife website.

Through automatic learning, like with rules of association and classification, georeferenced data can be used to predict and forecast common human movements, which are similar to a flock-like pattern (Krumm & Horvitz, 2007), and correlations between the places that people visit and users' needs can thus be established. The application of urban movement patterns has been widely discussed in recent studies (Vieira & Frías-Martínez, 2010; Zheng, Xie, & Ma, 2010), which may bring to light hidden and important information about the tendencies and interactions among people whose behavior is characterized by some very simple rules—such as do not bump into others, arrive at your destination, and maintain a safe following distance—that interact to generate intelligent group behavior.

By applying sophisticated techniques such as visual analytics and data mining in studies of human movement, different types of tendencies and patterns can be revealed, such as similarities among users' characteristics that establish behavior clusters. These could allow a company to conduct personalized marketing, and the State could also use this information to detect problems and make decisions about mobility and support in areas such as urban planning, public transport systems, and safety. This kind of knowledge could have prevented major problems such as that which occurred in China in the summer of 2010, when a traffic jam on the Beijing-Tibet freeway sprawled over 60 kilometers and lasted 9 days.

---

[18] Microsoft Research Asia. GeoLife, GPS Trajectories, 2013, http://research.microsoft.com/en-us/projects/geolife/ (GeoLife, 2013).

Feedback between real-time monitoring of traffic flow and computer models offers the proven ability to alleviate congestion problems. Furthermore, traffic modeling is suggesting new ways of planning roadway systems with better designs of road and intersection networks, along with intelligent management of traffic signals, rules, and the use of automated driving assistance systems with GPS technology. Optimal patterns of movement and behavior can be generated, increasing the flow of traffic, providing space for more cars, and reducing traffic jams, and thus improving road safety and minimizing the number of deaths. This also reduces pollution and improves a country's economy by significantly cutting the number of working hours lost and preventing thousands of millions of liters of fuel from being wasted on the world's roads.

Knowing how people move in relation to public spaces and how they interact among themselves allows leaders, architects, and urban planners to design cities more efficiently. Disasters in stadiums or concerts caused by crowd panic are thereby prevented by identifying bottlenecks, and isolated places prone to crime are detected.

Similarly, the data generated by the State can serve as a reference for the company or for industry. A specific example of this is the monitoring undertaken by different governments of Antarctic icebergs, formed by the separation of enormous blocks of ice from glacial platforms. Over the past 30 years, various research projects have studied the movement of icebergs in Antarctica using diverse technologies (Long, Ballantyne, & Bertoia, 2002; Ballantyne & Long, 2002). The National Ice Center (NIC) and the Brigham Young University Microwave Earth Remote Sensing Laboratory (BYU) have employed a wide variety of sensors and satellites to manually track large Antarctic icebergs and capture their positions.

The NIC and BYA have created a database for monitoring Antarctic icebergs, including the identification of icebergs between 1978, 1992, and 2009. On average, each ice floe is reported in 1 to 5 days. The high temporal sampling of the data set offers valuable information about ocean currents in the study area. This is extremely useful for sailors, providing them with more accurate positional information about the Antarctic region.

It is a well-known fact that the oceans play a key role in global warming (Arrigo & Dijken, 2002). Icebergs are vital in this, as they collect and cool water moving down from warmer latitudes, allowing variation and heat exchange between the ocean and the atmosphere; furthermore, they have side effects on the movement of currents. Monitoring groups of icebergs is therefore important for understanding the behavior and effects of these currents on global climatic patterns. In turn, it is in the State's interest to synthesize the information that is needed from the data sources to address the problem and create effective solutions.

The data generated by governments has been used by private companies as a reference for studies in another important area: ecological iceberg monitoring related to the fishing industry. Antarctic krill represents a multimillion-dollar industry registering more than 100,000 tons caught each year (Nicol & Foster, 2003). At the same time, the biological importance of krill in the Antarctic ecosystem is a subject of growing concern for conservation and surveillance. Monitoring and studies, however, are restricted by the size of the species. Initial research has indicated that the global distribution of krill coincides with the distribution of glaciers and ocean currents (Hewitt, 2003; Nicol, 2006). The discovery of flock-like patterns of movement in icebergs is clearly very important both for studies of global climate and marine biodiversity in Antarctica.

Situation 2. Conduct, hysteria, and market

The act of one person assaulting, humiliating, or harming another is unacceptable. For us, psychopathic behavior is aberrant; furthermore, it is outrageous when this behavior is shared by a community and accepted norms allow one person to hurt another. Thirty years ago, for example, racist and homophobic behavior was considered normal, but it is unacceptable in many countries today. Societies are unsustainable without shared norms, which are indispensable so as to avoid having to constantly make new decisions about how we should behave, dress, or speak. For this reason, we negotiate the balance between individual freedom and collective responsibility as a society.

This ability to transmit emotions and influence human behavior and decisions explains the genocidal conduct that spread like a disease under the regime of Nazi Germany (Enciclopedia del Holocausto, n.d.), when a sociopathic sentiment rose up against the race that the Nazis referred to as "impure." The subjects followed a leader (Adolf Hitler) and accepted rules without argument, as they believed he was on a higher plane of thought and action, and that all his decisions were therefore right. In this way they attributed complete moral responsibility to him, thus avoiding their guilt and accepting this behavior as correct.

Another example can be seen in democratic electoral processes, which tend to arise in an interactive manner among individuals, that is, in a negotiated manner whereby agents attempt to persuade others to make the same decision as them to reach a consensus. This complex phenomenon is very similar to that observed in microscopic models of magnetism, in which each magnetic atom in a glass pane influences those around it; in the same manner, many of our decisions are based on social interaction. Some agents are stronger than others, however, and some harder to influence, and that is precisely the reason why knowing a system in detail and understanding the networks of advice and persuasion, influence and consensus, and the laws they are determined by can make a great difference and tip the balance.

Another situation in which complex behavior is caused by agents interacting is crime and delinquency, these being related to the characteristics of the built environment where they occur. The majority of crimes can be considered a social interaction between criminal and victim, even if it is a counterproductive interaction. This is analogous to encounters of predators and their prey in nature; accordingly, location can have a major influence on the chances of being assaulted. This suggests that antisocial behavior is transmitted by example, and it can be seen that in public places that appear dilapidated and neglected, people will allow themselves to further damage the space.

Some of the world's police departments, such as the Los Angeles Police Department (Predpol, n.d.), are implementing crime observatories in which data is entered into intelligent algorithms that use statistical inference to predict places where future crimes may occur. The algorithm identifies patterns of vehicle theft and home burglaries using a similar approach to that used to predict the location of earthquake aftershocks. This approach can also be used by companies to understand market behavior. The applications for this type of system are limitless and help users anticipate events, but they also underline the importance of having access to relevant, timely, and accessible data at all times.

It is therefore paramount that tools be available to collect and process big data, such as data warehouses, data mining techniques, and visual analytics, to understand the collective psychological identity, its social norms, and its secret codes. The State, for example, is interested in knowing about episodes of spontaneous panic, mass hysteria, or phobias common in emergencies such as the threat of a natural disaster, in order to create suitable evacuation areas. The State could also use the principle of behavioral propagation to instill ethical and positive behavior, which would help establish a healthy culture of coexistence; likewise, it would be possible to detect criminal patterns to reduce crime. The company, meanwhile, is interested in identifying specific behaviors with the aim of segmenting markets and focusing efforts on detecting fashions or trends and, in general, creating strategies, preventative measures, and contingency and action plans to effectively meet the current needs of our complex society.

One crucial field that the university, company, and State have in common is the economy, as it is one of the pillars on which society is built. Economic behavior must be studied to understand social behavior; however, economic predictions are unreliable, as Amartya Kumar Sen[19] (Sen, 2011) stated: "Something is missing from conventional economic models that prevents them from describing, much less predicting, social behavior." Market crashes and economic crises occur, as the economy cannot operate in an appropriate way because of disturbances that originate outside the system; in other words, the global economy should be recognized as an incredibly complex system, and this science should therefore make use of previously produced knowledge about complexity from other sciences such as biology or physics.

Accordingly, economic analysis requires massive data input and countless types of variables such as transactions, financial activities, interaction networks, and risk studies, along with a sociocultural context and the combination of information generated by the State to plan possible future situations and foresee economic crises or depressions.

Eurace[20] is a model created in Europe based on millions of agents that simulate a fictitious economy, including markets for labor, goods, credit, and finance. In addition, the model has a spatial element: the companies and workers are located in real places, linked through social and business networks. The objective is to use this model as a testing ground for aspects of European economic policy and to implement new, previously untried theories that are based on complexity science and that influence the behavior of economic and financial markets.

Conclusion

At present, the majority of models are still too abstract to provide reliable conclusions about the real world and establish complex behavior that can emerge from simple rules. The university, company, and State must therefore make a concerted effort to develop efficient tools and strategies for storing, sharing, consolidating, and analyzing big data and the information it contains in a comprehensive and synergistic way to create models of knowledge.

The challenge is to understand the complexity of our societies based on the data that has been generated up to this point, to determine how, when, and why the collective and complex behavior of our society changes, and thereby make the best decisions for the common good.

---

19 Bengali (India) philosopher and economist awarded the Nobel Prize in Economics in 1998.

20 http://www.eurace.org/

References

Arrigo, K. & Dijken, G. van. (2002). Ecological impact of a large Antarctic iceberg. *Geophysical Research, 29*(7), 6-9.

Ball, P. (2012). *Why Society is a complex Matter*. New York: Springer.

Ballantyne, J. & Long, D. (2002). A multidecadal study of the number of Antarctic icebergs using scatterometer data. *Geoscience and Remote Sensing, 00* (C), 3029-3031.

Butz, A., Fisher, B., Christie, M., Krüger, A., Olivier, P., & Therón, R. (2009). *Smart Graphics*. Salamanca: Springer. Retrieved from http://www.amazon.ca/Smart-Graphics-International-Symposium-Proceedings/dp/364202114X

Dettki, H., Ericsson, G., & Edenius, L. (2004). Real-time moose tracking: An internet based mapping application using GPS/GSM-collars in Sweden. *Alces*, 40. Retrieved from http://bolt.lakeheadu.ca/~alceswww/Vol40/Alces40_13.pdf

El País. (2011, February 10). *315 veces más información que granos de arena*. Retrieved September 16, 2013, from http://sociedad.elpais.com/sociedad/2011/02/10/ actualidad/1297292421_850215.html

El País. (2012, April 24). *Facebook camina hacia los mil millones de clientes*. Retrieved from http://tecnologia.elpais.com/tecnologia/2012/04/24/actualidad/1335257063_449543.html

EMC. (n.d.). *EMC*. Retrieved September 16, 2013, from http://colombia.emc.com/ campaign/bigdata/index.htm

Enciclopedia del Holocausto. (n.d.). *Enciclopedia del Holocausto*. Retrieved 16 September, 2013, from http://www.ushmm.org/wlc/es/article.php?ModuleId=10007792

Frank, A. U., Raper, J., & Cheylan, J. P. (2001). *Life and Motion of Socio-Economic Units*. London: Tylor & Francis.

Gache, F. L. & Otero, D. (2010, November 30). Adam Smith: la mano invisible o la confianza. *Visión de Futuro, 14*(2), 28.

Gantz, B. J., Reinsel, D., & Shadows, B. D. (2012). *The Digital Universe in 2020: Big Data, Bigger Digital Shadows and Biggest Growth in the Far East Executive Summary: A Universe of Opportunities and Challenges*, 2007. Retrieved from: http://idcdocserv.com/1414

García, J. C., Sánchez, M. M., & García, M. N. (2006). *Multiclasificadores: métodos y arquitecturas*. Salamanca: Universidad de Salamanca.

GeoLife. (2013). *Geolife*. Retrieved September 16, 2013, from http://research. microsoft.com/en-us/projects/geolife/

Gudmundsson, J. (2004). *Efficient detection of motion patterns in spatio-temporal data sets*. Proceedings of the 12th. Retrieved from http://dl.acm.org/citation.cfm?id=1032259

Heredia, D. P. (n.d.). *Del bit a la revolución informática*. Retrieved September 16, 2013, from http://www.matematicasyfilosofiaenelaula.info/conferencias/Shannon.pdf

Hewitt, R. (2003). An 8-year cycle in krill biomass density inferred from acoustic surveys conducted in the vicinity of the South Shetland Islands during the austral summers of

1991-1992 through 2001-2002. *Aquatic Living Resources, 16*(3), 205-213. DOI: 10.1016/S0990-7440(03)00019-6.

Iwase, S. & Saito, H. (2002). *Tracking soccer player using multiple views. Proceedings of the IAPR Workshop on Machine*, 1-4.

Krumm, J. & Horvitz, E. (2007). Predestination: Where Do You Want to Go Today? *Computer, 40*(4), 105-107. DOI: 10.1109/mc.2007.141.

Long, D. G., Ballantyne, J., & Bertoia, C. (2002). Is the Number of Icebergs Around Antarctica Really Increasing? *Eos, Transactions American Geophysical Union, 83*(42), 469-474.

López, J. M. & Herrero, J. G. (2006). *Técnicas de análisis de datos*. Madrid: Universidad Carlos III.

Makris, D. & Ellis, T. (2002). Path detection in video surveillance. *Image and Vision Computing, 20*(12), 895-903. DOI: 10.1016/S0262-8856(02)00098-7.

Maldonado, C. E. & Gómez Cruz, N. A. (2010). *El mundo de las ciencias de la complejidad. Un estado del arte*. Universidad del Rosario. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1929020

Nicol, S. (2006). Krill, Currents, and Sea Ice: Euphausia superba and Its Changing Environment. *BioScience, 56*(2), 111-111. DOI: 10.1641/0006-3568(2006)056 [0111:kcasie]2.0.CO2.

Nicol, S. & Foster, J. (2003). Recent trends in the fishery for Antarctic krill. *Aquatic Living Resources, 16*, 42-45.

Pereira, R. T.; Millán, M., & Machuca, M. (2006). New Algebraic Operators and SQL Primitives for Mining Classification Rules. *Computational Intelligence* San Francisco, California, United States.

Predpol. (n.d.). *Predpol*. Retrieved September 16, 2013, from http://www.predpol. com/

Reynoso, C. (2007). *Edgar Morin y la complejidad: elementos para una crítica*. Buenos Aires: Grupo Antropocaos.

Romero, A. O. (2011, March). *Mining moving flock patterns in large spatio-temporal datasets using a frequent pattern mining approach*. Retrieved from http://tomx.inf.elte.hu/twiki/pub/Tudas_Labor/2012Summer/thesis_colfuturo.pdf

Rúa, M. (2012, July 15). *El nuevo oro se llama big data*. Retrieved September 16, 2013, from http://www.lanacion.com.ar/1490304-el-nuevo-oro-se-llama-big-data

Salazar, B. (2007). Thomas C. Schelling: la paradoja de un economista errante. *Revista de Economía Institucional, 9*(17), 131-152.

Schelling, T. (1981). *Strategy of conflict.* United States: Harvard University Press.

Schelling, T. (2008). *Arms and Influence.* United States: Yale University Press.

Schmarzo, B. (2011). Análisis de Big Data. *EMC, 8*.

Sen, A. K. (2011). *Desarrollo y crisis global.* Madrid: Universidad Complutense de Madrid.

Stewart, I. (2007). *¿Juega Dios a los datos?* Barcelona: Crítica.

Vieira, M. & Frías-Martínez, E. (2010). *Querying spatio-temporal patterns in mobile phone-call databases. Mobile Data.* Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber

Vieira, M. R., Bakalov, P., & Tsotras, V. J. (2009). On-line discovery of flock patterns in spatio-temporal data. *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems – GIS*, 09, 286–286. DOI: 10.1145/1653771.1653812.

Zambrano Reyna, R. (2011). Impacto de internet y las redes sociales en la conducta humana. Retrieved from http://www.cec.espol.edu.ec/blog/wp-content/uploads/2012/08/IMPACTO- DE-INTERNET-Y.pdf

Zheng, Y., Xie, X., & Ma, W. (2010). GeoLife: A collaborative social networking service among user, location and trajectory. *IEEE Data Engineering Bulletin, 33*(49), 32-40. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.165.4216&rep=rep1&type=pdf

Zheng,Y.,Zhang,L.,Xie,X., & Ma,W.(2009).Mining interesting locations and travel sequences from GPS trajectories. *Proceedings of the 18th International.* (49). Retrieved from http://dl.acm.org/citation.cfm?id=152681

Zoya, L. G. & Aguirre, J. L. (2011). Teorías de la complejidad y ciencias sociales. Nómadas. *Revista Crítica de Ciencias Sociales y Jurídicas, 30*(2), 147-166.