Maria Begoña Peña Lang

Big Data STEM Education: "The Skills Key"

University of the Basque Country

Let's talk about the socio-economic context in which the European Union is currently immersed, which is undoubtedly characterized by an uncontrollable and vertiginous tendency towards the globalization of information and knowledge. The role of Education becomes relevant in a way that is inevitable, there's no way to go around it. We believe that Education must become the true vanguard beacon of the deep transformation process that our societies are experiencing. And given the absence of discontinuity that can be deduced from this great social change process, it is therefore necessary that the concept of Education also acquires and maintains a continuity, for which we must first overcome the gaps or discontinuities that occur throughout the formation of the individual throughout its life path.

We are talking about overcoming the stagnation that occurs between the different formative periods through which a person goes through; a rupture from isolation and an aperture to continuity that has come to be known as *lifelong learning*. Though it is true that this concept is not a novelty in itself, since in ancient classical texts we find explicit references to the need to extend education uninterruptedly throughout the life of a person, it is also true that currently the global market to which we are subject imposes with force the urgency to revitalize this concept, but with a new approach that adds value to Education.

Now, the implementation and strengthening of the concept of *lifelong learning* entails a series of challenges that must be faced, the first of which is the aperture to new and multiple pedagogical dimensions that enable the individual to acquire basic skills and abilities in a truly flexible way, from the first education levels to professional training. And inevitably linked to this first challenge we find the full desirability of fostering social ties that accommodate diversity and equity.

Lifelong learning also should become a way of blurring the socio-economic discontinuities and knowledge gaps between different social groups. The challenge, therefore, is to create a new pedagogical foundation, theoretical in nature, but with a practical application that must be reflected in the development of curricular proposals that respond to the demands of the great heterogeneous social groups.

In this sense, scientific knowledge, put at the service of the learning process of young people and adults, in its broadest sense, puts us before a second challenge; how to establish, in a consistent manner, all the modalities of certification of knowledge and competences, and to do so on a scale that exceeds the parameters and contingencies of each country. Europe is the context within which we have to design, institutionalize and evaluate the mechanisms that allow integrating this multiplicity of educational offerings. This implies a complex network of interinstitutional agreements on the different values of different types of knowledge, in accordance with the common objectives pursued, without in any way avoiding the matter of financing of such international consensus.

Consequently, considering all of the above, it seems evident that the greatest challenge lies in involving all actors possible in the educational process, including schools and colleges, universities, national and international institutions, and the business world. Only this way will the rupture mentioned above be possible. And this context is in turn the framework in which the main objectives that the European Commission has designated for *lifelong learning* are contained:

«-Improve and increase the mobility of pupils and staff across the EU.

-Enhance and increase school partnerships across the EU.
-Encourage language learning, ICT for education, and better teaching techniques.
-Enhance the quality and European dimension of teacher training.
-Improve approaches to teaching and school management. The term "STEM education" refers to teaching and learning in the fields of science, technology, engineering, and mathematics. It typically includes educational activities across all grade levels—from pre-school to post-doctorate—in both formal (e.g., classrooms) and informal (e.g., afterschool programs) settings».

The last of these objectives puts us before the evidence of the important role that STEM studies should play, a term that refers to cross-curricular teaching and learning in the fields of science, technology, engineering and mathematics. An STEM-based education aims to offer opportunities to students so they can do meaningful and coherent work, while at the same time positioning them as learners of their own learning, helping them understand how knowledge is grounded and developed, and to solve world problems Through experimental learning projects with STEM, connecting scientific areas, engineering and mathematics, and using technology as an authentic tool for integrating information and knowledge, it is created a non-discontinuous space for discovering innovative solutions and connections between the school, university and the company; a path to address the deficiencies referred to at the beginning of this lecture.

That's why we believe that implementing a STEM education will pave the way for new teaching practices and strategies that will create an optimal environment for students to experiment, discover, design, create, construct and review; something which would be equivalent to receiving optimal preparation for the jobs of the future. The competences sought to be strengthened through STEM are not chosen by chance or merely by a whim, but they rather obey those principles that will allow young people to be competitive in an increasingly globalized economy. These are skills or abilities that are increasingly necessary to live in an eminently technological world, determined by ever evolving telecommunications and a knowledge-based economy. Today we already have strong evidence that this is the path to follow; we have indicators that clearly show that any professional, regardless of the academic field from which he or she comes, must have developed these competencies to effectively perform their work.

In sum, the skills that the 21st century imposes on each individual must be well deployed and strengthened, as they require an orientation that corresponds perfectly with STEM programs; programs that, through the most varied projects, incorporate group and cooperative work, which enhance the most practical aspects of knowledge, and contribute to a good research profile. Innovation, cooperation and the resolution of complex problems, as well as the capacity for critical analysis, are those skills that to a great extent, guarantee professional success and preparation to educate citizens that are capable of making decisions on issues of social importance and diverse, such as health, environmental quality, energy efficiency, national security or the use of natural resources.

Finally, we need to emphasize the importance of scientific and technological innovation for the sustained growth of developed countries, as well as design strategies that give opportunities to those who are in the process of development to reduce their dependence on the so-called "First world", thus increasing their competitiveness in the world market and improving the living standard of its citizens. Rodger W. Bybee summarizes the concept of STEM that we have characterized in the following terms:

«*Most, it means only science and mathematics, even though the products of technology and engineering have so greatly influenced everyday life. A true STEM education should increase students understanding of how things work and improve their use of* *technologies. STEM education should also introduce more engineering during precollege education. Engineering is directly involved in problem solving and innovation, two themes with high priorities on every country's agenda. Given its economic importance to society, students should learn about engineering and develop some of the skills and abilities associated with the design process*».

From what has been stated above, two essential ideas can be extracted to finally consolidate a *lifelong learning* concept that exactly corresponds to the parameters and objectives that we have been pointing out: 1) to have powerful tools to do a basic empirical study that allows us to obtain and interpret the data needed to break the gaps between the different formative stages of a person, and 2) to create a conglomerate of multiple and heterogeneous institutions that work as partners and seek educational innovation.

Regarding the first indicated idea, we think that *Big Data* and *Mining Data* are the two tools that will facilitate the enormous collection and interpretation of data that is required. *Big Data* is the term usually used to refer to such an enormous amount of data that exceeds the capacity of conventional software to be captured, classified and processed within a reasonable amount of time. It is a fact that in recent years the volume of data has grown exponentially, since it is necessary to have instruments to process and interpret them effectively and quickly. In the field of Education this ability to capture and interpret would greatly favor the decision making process that would be directed towards the goal of *lifelong learning*.

As for *Mining Data*, the *Knowledge Discovery in Databases* (or KDD) analysis stage, as a specific field of statistics and in general of the so-called computer sciences, can show us its validity when it comes to discovering common patterns on large volumes of data sets, something which is undoubtedly essential to our purposes. We need systematic data compilation and its subsequent classification, but also an adequate interpretation that will lead us to the determination of models that can be used for prediction. Thus, *Mining Data*, based on artificial intelligence methods, automatic learning and statistics, also serves our purposes in the field of Education, since its function is precisely to extract information from a set of data and transform it into a understandable structure for later use. In this sense, in regards to the use of Mining Data in Education, Jiménez Galindo and Álvarez García, from Universidad Carlos III of Madrid argue that:

"Educational data mining offers many advantages by comparing it with more traditional research paradigms related to education, such as laboratory experiments, sociological studies, or design research. In particular, the creation of public repositories of educational data has created a foundation that makes possible the mining of educational data. More specifically, the data from these repositories are fully valid (since they are actual data measured from the performance and learning processes of real students in educational settings, taken from learning tasks), and increasingly more accessible to begin an investigation. These points allow researchers to save a lot of time on tasks such as finding people (such as schools, teachers and students), organizing studies and collecting data, since they are directly accessible "[1].

Certainly this term is in vogue, though it is often misused or inaccurately used to refer to any form of large-scale data or all information processing (which would include collection, storage, analysis and subsequent statistics metrics), mixing with it any type of computer support system, artificial

---

[1] JIMÉNEZ GALINDO, Álvaro y ÁLVAREZ GARCÍA, Hugo (2010); *Minería de Datos en la Educación,* Universidad Carlos III de Madrid.

intelligence, automatic learning and even business intelligence. However, the key for the subject that concerns us is the term *discovery*, usually defined as 'the detection of something new'. According to this, the priority task of data mining becomes the automatic analysis of large amounts of data, from which interesting patterns that were previously unknown could be extracted, such as groups of data records (cluster analysis), unusual records (detection of anomalies), or as dependencies (mining by association rules).

At this point, such patterns can be considered as a sort of summary of all the data being collected. Subsequently, an additional analysis begins (automatic learning and predictive analysis), which places us in front of a series of complex systems. By definition, every *complex system* is formed of various parts that are interconnected or interlaced, whose links provide new information that was not previously visible to the observer. The result of these interactions between elements is the emergence of new properties that are not possible to explain in an isolated way. Thus we call them emergent properties. Complex systems are opposed to other systems that though they are also constituted by different parts, the relations established between them don't add additional information.

In this other type of systems, it is sufficient to know how each of them works to understand them as a whole. In a complex system, on the contrary, hidden variables appear whose knowledge makes it possible to advance in a more precise understanding and interpretation of the data that have been collected. In short, a complex system provides us with information that goes beyond the simple isolated understanding of each component, so that to describe it, we need to know the operation of the parts and also understand the functioning of the system as a whole, as their parts are related to each other. This is the context that can enrich the most the application of these techniques to the field of Education. If we want to move towards a *lifelong learning* that is consistent with the capacity to evolve the educational landscape, we can't neglect these valuable tools.

In conclusion, the *lifelong learning* process to which we aspire to implement must combine the interaction of all parts involved in the formation of a person, in all its stages and in a continuous way, with the tools that the new technologies put within our reach. In other words, we seek an effective cooperation of the most diverse educational agents, supported by empirical research that draws from data, patterns and complex systems that can be extracted and configured from the application of *Big Data* and *Mining Data*.