

Learning Education: An 'Educational Big Data' approach for monitoring, steering and assessment of the process of continuous improvement of education

Paper presented in

European Conference in the Applications of Enabling Technologies, 20-21 November 2014, Glasgow, Scotland

Wouter Vollenbroek¹, Knut Jägersberg², Sjoerd de Vries², Efthymios Constantinides¹

¹University of Twente, Enschede, The Netherlands

²NHL Hogeschool, Leeuwarden, The Netherlands

w.vollenbroek@utwente.nl

k.jagersberg@alumnus.utwente.nl

sjoerd.devries@nhl.nl

e.constantinides@utwente.nl

Abstract: Changing regulations, pedagogy and didactics worldwide, have ensured that the educational system has changed severely. But the entrance of Web 2.0 and other technologies had a significant impact on the way we educate and assess our education too. The Web 2.0 applications also increase the cooperation between stakeholders in education and has led to the phenomenon 'Learning Education'. Learning Education is a term we use for the phenomenon where educational stakeholders (i.e. teachers, students, policy-makers, partners etc.) can learn from each other in order to ultimately improve education. The developments within the Interactive Internet (Web 2.0) enabled the development of innovative and sophisticated strategies for monitoring, steering and assessing the 'learning of education'. These developments give teachers possibilities to enhance their education with digital applications, but also to monitor, steer and assess their own behavior. This process can be done with multiple sources, for example questionnaires, interviews, panel research, but also the more innovative sources like big social data and network interactions. In this article we use the term 'educational big data' for these sources and use it for monitoring, steering and assessing the developments within education, according to the Plan, Do, Check, Act principle (PDCA). We specifically analyze the Check-phase and describe it with the Learning Education Check Framework (LECF). We operationalize the LECF with a Learning Education Check System (LECS), which is capable to guide itself and change those directions as well in response to changing ways and trends in education and their practices. The system supports the data-driven decision making process within the learning education processes. So, in this article we work on the LECF and propose and describe a paper-based concept of the – by educational big data driven – LECS. Besides that, we show the possibilities, reliability and validity for measuring the 'Educational Big Data' within an educational setting.

Keywords: Learning Education, Learning Education Check Framework, Educational Big Data, Web 2.0, PDCA

1. Introduction

The last decades brought the linear and mechanistic worldview to an end and heralded the conception of an ecological universe – a holistic, dynamic, and inextricably connected system in which everything seems to affect everything else (Hesselbein, Goldsmith & Beckhard, 1997). The current networked society provides many tools and possibilities to individuals but also to organizations, networks and the society, to engage in continuous learning (Pang, Jie & Xu, 2013). A development that is stimulated by the options offered by Web2.0 applications and outcomes (i.e. West, 2012). An effective organization, which can be an educational institution, is according to Slater & Narver (1995) an organization providing configurations of management practices that facilitate the development of the knowledge as the basis for competitive advantage. Pedler et. al (1991) define such an organization as a learning organization following a learning perspective: "an organization that facilitates the learning of all of its members and continuously transforms itself in order to meet its strategic goals". But the initial idea of a learning organization has been coined by Senge (1990) and defined as: "the encouragement of organizations to shift to a more interconnected way of thinking". The definitions of Senge (1990) and Pedler et.al. (1991) cover a significant part of learning education; we use the core of this definition to define learning education as: *"The learning process where all interconnected members learn of each other within open educational networks and continuously transform themselves in order to meet its strategic goals by using the latest technological developments in the field"*. This process of lifelong learning forms the backbone of learning education.

In learning education all educational stakeholders such as teachers, students, policymakers, and educational researchers learn from each other in order to continuously improve their current practices and ultimately the education. Many educational institutions nowadays pursue the active educators' engagement in innovative processes using on- and offline tools allowing them to improve educational

practices; such engagement is facilitated by online and offline social platforms where educators interact with other educators. These social platforms generates a lot of data, that is already online (big social data and learning analytics) or can be published online (questionnaires, interviews and panel research). So the current stage of the Internet technology i.e. the Interactive Web or Web 2.0 is one of the game changers in learning (Brown & Adler, 2008) since it enables learning education, provides open education and provides innovative and sophisticated strategies for monitoring, steering and assessing this process. In this paper we present the Learning Education Check Framework (LECF) based on relevant educational data analysis. The LECF underpins the Learning Education Check System (LECS), that facilitates the collection, analysis and visualization of relevant educational data, defined as educational big data.

2. Literature review

The learning education check system is to support a self-sustaining mechanism to all stakeholders of education. LECF makes use of multiple resources and methods, the more traditional ones like questionnaires and interviews and the more progressive ones like big data. The so-called big data makes it possible to mine learning information for insights regarding performance and learning approaches (West, 2012). We need to design a system, which is capable to guide itself and change those directions as well in response to changing ways and trends in education. A system that supports a data-driven decision making process within education. According to Marsh, Pane and Hamilton (2006) the data-driven culture that arises within educational institutions refers to teachers, principals, and administrators that systematically collect and analyze different types of data, including input, process, outcome and satisfaction data in order to guide a range of decisions to help improve the success of students, teachers and schools. This is why we also seek inspiration in cybernetic principles / control theory (Wiener, 1948), as learning education can only be ideated as a self-regulating system if to be maintained sustainably. We build the learning education check system to exert cybernetic control over itself: Factually all stakeholders in education influence each other by these principles, so we take a psychological perspective upon control theory, as done by i.e. Carver & Scheier (1982, 1999).

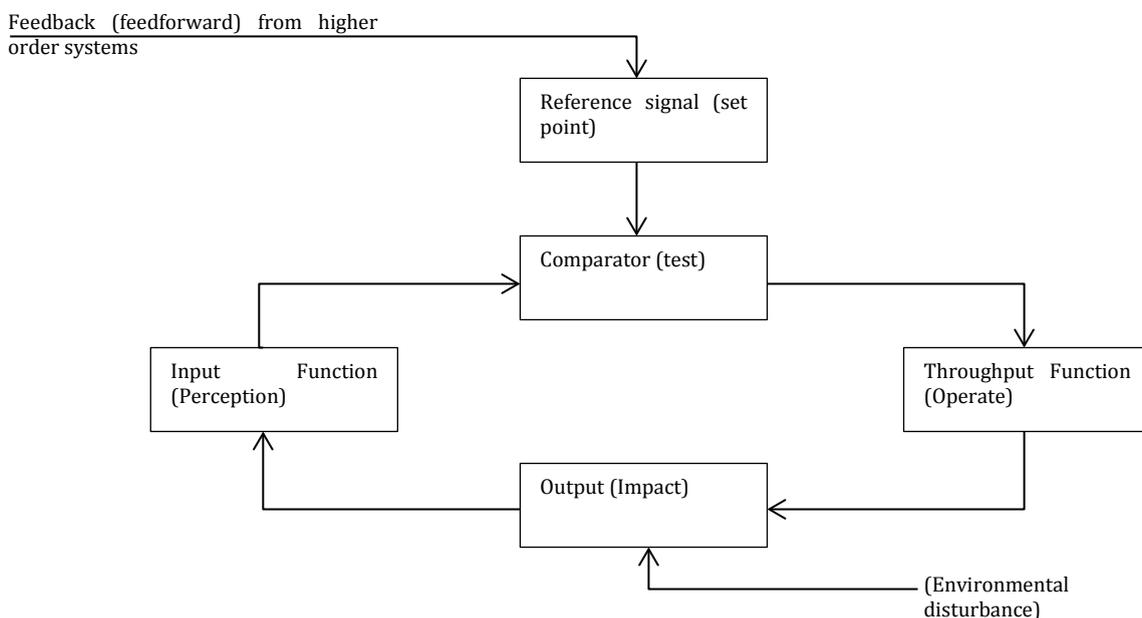


Figure 1: A Test-Operate-Test-Exit or TOTE unit (adapted from Carver & Scheier, 1981).

The basic principles of the self-regulatory mechanisms which we had to take into account whilst designing our learning education framework are quite simple:

1. All stakeholders of education need to get appropriate information about their environment about the state of affairs in education (input function in cybernetic terms)
2. They need to co-control the direction of the educational system, meaning they have to give goals to the governance processes of education (reference signal / set point)
3. Attention of the stakeholders towards these goals must be warranted so they compare the current state of affairs in education with whatever has been found as 'good' education
4. Engagement towards action must be supported by a (digital) environment, sustaining the motivation to reach those goals
5. The environmental change must be electronically stored (i.e. data warehouse) in a way to create a 'physical' basis for perceiving new changes of the environment in the next round (output / impact,

which can also monitor environmental disturbances to the education processes)

This negative information feedback loop regulates education towards the co-controlled and co-created goals of education in a society. The smartness of a learning education check system comes from an invisible secondary feedback loop *system* (feedback from multiple if not infinite higher order systems) to determine future directions. It must be conceived in a way not to be conceived in a strictly fixed way so it saves the dynamic momentum necessary for sustainable guidance of education, as patterns of innovation (i.e. in education) have been suggested to have strong similarities with patterns of deterministic chaos (Radzicki, 1990). Therefore, the input list of our learning education check system is always only a temporary depiction and crude picture of information flows of main influence to guide the education, it is never finished or complete. In a way, the whole system must always remain capable of re-structuring and reorganizing itself, even dramatically to govern education in a sustainable way.

The key word to make the second order feedback loop infinitely encompassing of societal goals is transparency. By exploiting the technological advances offering more insights in society issues by learning from big data analysis, we can now create a learning education check system, which is close to being informed by all possible information flows or reference signals (or factually higher order feedback loops) of a society. The process, which makes this possible is called open innovation. Build on these insights, we have designed a learning education check system where all stakeholders of education could influence each other in our networked society.

However, the introduction of Web 2.0 applications in educational systems has also led to new challenges. The SLEPT-model, a model that has its origins in organizational science, identifies fields where such challenges arise; the model is meant to make businesses aware of changes in their external environment in order to adjust their strategy and remain competitive in their markets (Doole & Lowe, 2004). The external environment can be described by five fields: Social, Legal, Economics, Politics and Technology. These five fields are not only relevant in business environments, but are useful in other domains as well, one of them being the educational learning networks. In such networks on the *social* level, many participants are for example experiencing an 'information overload' (Kop, 2011), since the production and dissemination of information is free and regulation of information streams is largely the responsibility of the members of a network. With the introduction of Web 2.0 applications network members are sharing vast quantities of information, which makes it often difficult for people to identify the information relevant to them.. The 21st Century Skills are created to help them in this new situation by turning users to digitally literate users. Everyone should know what information is relevant and valuable to his or her own knowledge domain. Another challenge arising with the Web 2.0 environments in knowledge sharing platforms is of a *legal* nature (Franklin & Harmelen, 2007). This refers to a trend to increasingly use online material improperly, without for example referencing to the original authors or sources. This way internet users undermine existing copyrights or intellectual property. The creative commons licensing facilitates knowledge distribution, without complicated copyright restrictions (Carroll, 2006). On the *economic* level, the integration of Web 2.0 in the online world increasingly forces organizations to adopt innovative and adaptive business models (Teece, 2010). Examples of such business models are the freemium and premium approaches where users use the network for free (freemium), but pay when more functionalities or services are added. (premium). On the *political* level, the ownership of knowledge is a point of discussion (Raban & Rafaeli, 2006). According to Raban & Rafaeli (2006) ownership makes a difference, it serves to increase sharing of information. But good agreements and guidelines as to the ownership regulations should be made. Finally on the *technological* level one of the challenges that exists with the rise of Web 2.0 applications is the connection between Web 2.0 applications in order to develop web services (Pullen et al, 2004), for example between a Learning Management System and an external website. A reliable and acceptable integration and connection between two systems requires a lot of knowledge and work of web developers. In the development of learning education networks, it is important to take these challenges in consideration.

3. Learning Education Check Framework

The LECF is a process where individuals learn from each other in order to improve their knowledge, competencies, practices and experiences. To support the development and distribution of knowledge, Web 2.0 applications are nowadays almost indispensable (Brown & Adler, 2008). To describe these processes which occur on- and offline, we use the Plan-Do-Check-Act cycle of Deming (1950); the LECF is a derivative of the Plan-Do-Check-Act cycle (PDCA-cycle).

- The Plan phase establishes the objectives and processes necessary to deliver results in accordance with the targets or goals;
- In the Do phase, the established plan is implemented in a certain setting;
- In the Check phase the actual results are analyzed; this is the phase we examine in this article;
- In the Act phase corrective actions based on what resulted from the check phase are taken.

If the change does not produce the desirable results the cycle starts again with another plan. The

implementation of the PDCA-cycle will be explained by means of an example: The process starts with the decision of the educational institutions to promote the interdisciplinary learning of teachers i.e. become a learning school (Plan). To implement this plan, the educational institution can encourage teachers to do a practical study (Do). The results of the study can be collected from many types of data and with various methods, for example from big (social) data, metadata, learning analytics, using questionnaires, interviews and panel research. In order to use this data in a valuable way, it is according to Marsh, Pane & Hamilton (2006) recommended to analyze the data in order to monitor, steer and assess the developments in the practical study and make a data-driven decision based on the analysis (Check). The recommendations from the check-phase are put into practice in order to fulfill the ultimate plan (Act).

In this article we describe the 'Check-phase' in more detail, this phase is described with the Learning Education Check Framework. To continually monitor and evaluate the movements and developments within learning education, we introduce an infrastructure for the continuous monitoring of these developments in the form of the Learning Education Check System (LECS). A LECS can be compared with a Learning Management Systems, where learning analytics – defined as the area of research and practice that aims to describe and evaluate the definition, collection, analysis and use of data resulting from technology use (Haythornthwaite, De Laat, Dawson & Suthers, 2013) – are used to describe the learning process of students (West, 2012). However, LECS not only use the general learning analytics available within Course & Learning Management Systems, but extends this with other sources such as big social data, questionnaires, interviews, panel research and a network monitor. It uses the Learning Education Check Framework to work in a systematically way on the analysis of relevant data. Figure 2 shows the components of our infrastructure, we have several complementary sources to feed our database, which supports a dashboard-supported representation of the 'Check-phase'.

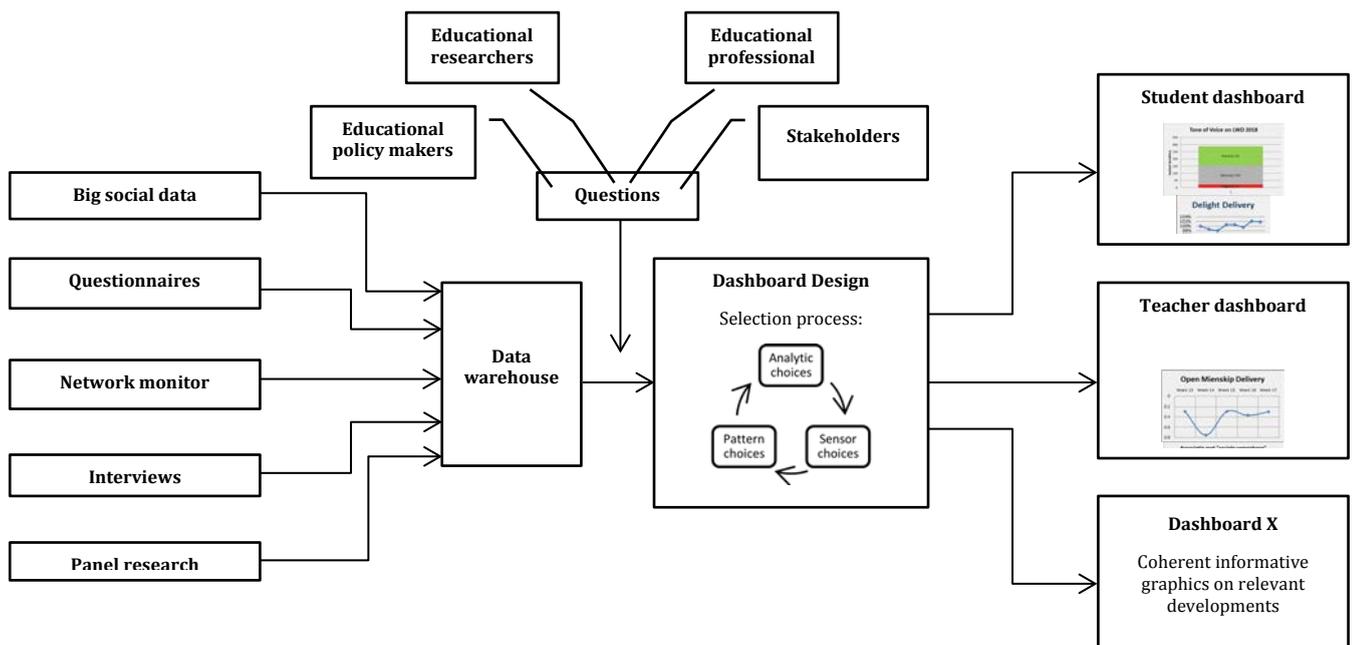


Figure 2: Learning Education Check Framework

The private data available to certain educational stakeholders is collected by means of questionnaires, learning analytics and interviews. This data can be measured and analyzed with traditional methods of research. The unique part of this research article is the development of a data warehouse and related analysis methods where all types of data can be collected, analyzed and visualized. We describe the collection of all relevant data for a certain context in an online environment as 'educational big data'. On the left side of Figure 2, we have big social data, that consists of public and private data from Web2.0 resources such as social media, Learning Management Systems and Course Management Systems. Besides that, the infrastructure can handle data from questionnaires in a systematic manner. Next to this, we have also a modular network monitor (questionnaire) and a wide array of computer supported social network analysis methodologies, which we can use to analyze the data. The integration of interviews can deepen the content of the platform. The last resource is the panel research, the structured data from panel research can be collected and analyzed within the system and can be connected to other resources. LECS is only intended to use data that is publicly available and private data that is only available for certain educational stakeholders, such as teachers, students, policymakers, researchers et cetera and should be only available to teachers and their principals. The educational big data can be collected with conventional and non-conventional data collection applications (i.e. data-scrapers, online survey software, etc.) and be stored in a data warehouse.

Educational policy makers, educational resources, educational professionals and other educational stakeholders can request for learning analytics (i.e. popularity of a project) and use the data warehouse to select the data relevant for the construction of learning analytics. The learning analytics can be constructed by connecting sensors (i.e. logs or mentions) to each other. An example of these learning analytics is the educational reputation of an actor or the educational influence of a teacher. When the data has been stored, it can be analyzed in a dashboard where analytical choices should be made, sensors should be chosen and patterns can be identified (central part of Figure 2). The last step in the framework is to visualize the analyses in an audience-centered dashboard (right part of Figure 2).

One of the unique parts of educational big data, is the big data component. Zikopoulos & Eaton (2011) describes big data with four characteristics: volume (amount of data), variety (diversity of data), velocity (speed of data acquisition) and veracity (reliability of data). The essence of big social data is comparable with big data, but differs in the content. Where big data contains all structured and unstructured data that derives from all kinds of digital applications, is the term big social data mainly associated with the data that is produced within the social media. Cambria, Rajagopal, Olsher & Das (2014) points that big social data contains sources such as social network analysis, multimedia management, social media analytics, trend discovery and opinion mining.

In this article we describe the process from educational big data sensor to visualization within the LECS. The LECS is an independent operating system that can be connected with other higher education systems (CMS and LMS) with Learning Tools Interoperability (LTI) connections (García-Peñalvo, Conde, Alier, & Casany (2011). The basic idea of LTI is to establish a standardized way of integrating rich learning applications (i.e. LECS) with other interactive education platforms like learning management systems, portals, or other environments. This process is divided into the following three steps: 1) educational big data collection (scraping), 2) educational big data analysis and 3) educational big data visualization. We will describe these three steps briefly and give some examples of relevant software-application to conduct the analysis of one of the essential components of the learning education check system: media sensor network.

3.1. Educational Big data Collection

We identified four educational big data collection methods that use quantitative and qualitative sensors from data already available by teachers, saved in a data warehouse. This data – mentioned as media sensor network - comes from online resources such as big social data, metadata and learning analytics. The media sensors within the media sensor network are the measurements and tools, which characterize what is going on in digital media, in a broad sense of that word. A sensor network comprises sensors (like a construct in a questionnaire), but most are integrating information from several meters (like an indicator in a questionnaire for a particular construct). The learning analytics describe the development of a learning process, the data to monitor this development derives from different sources. Some of these sources are: media sensor network from i.e. Learning Management Systems, social media etc., questionnaires, network monitor from i.e. social media, interviews and panel research. But to collect the big social data, different data-scrapers can be used (see Appendix 1):

- Simple educational big data analytics describe the basic characteristics of certain educational big data use. Examples of these analytics are: popularity, activity or audience knowledge.
- Educational big data content analytics describe the content aspects of certain educational big data use. Examples of these analytics are: opinion mining, sentiments and genre development.
- Educational big data network analytics describes the network characteristics of certain educational big data use. Examples of these analytics are: interaction overview and likes and comments.
- Cross media analytics describes the relation between certain educational big data use and aspects of other media such as newspapers and TV. One of the examples of these analytics is the cross reference.

3.2. Educational Big data Analytics

The educational big data collection mainly consists of simple sensors (i.e. number of posts or comments), these sensors are analyzed in combination with data analysis software-applications. The successful analysis of these educational big data sensors depends mainly on the way of editing and processing the data. It chasing through three stages: analytic choices, sensor choices and pattern choices. One can for example want to identify the social reputation of a teacher (analytic), for that reason the popularity and activity of one should be connected to each other (sensors) and ultimately a pattern could be identified or created to strengthen the analytics.

3.3. Educational Big data Visualization

Data visualization plays a major role in the educational big data research process, since we nowadays not only use it for the presentation of results, but also for the analysis of data (McAfee, Brynjolfsson,

Davenport, Patil & Barton, 2012; Mazza & Dimitrova, 2004). For example, big data visualization can be used to get insight in the distribution of a sensor and to show relations between actors, interactions et cetera (social network analysis). In the LECS the offer to show the analysis with a visual representation, is an important component. The users of LECS need to gain in a clear and visually appealing way insight into these analyzes, so there will not arise any ambiguity about the interpretations. For that reason, within LECS adaptive dashboards can be developed based on the context and situation.

4. Conclusions and discussion

The Learning Education Check System, described in this article, is a paper-based concept. In the near future this concept will be further developed in order to contribute to contemporary education and to provide guidance on the monitoring and control of it. The core of the educational big data analysis in the Learning Education Check Framework is to show the progress that can be made with the learning education principle, for teachers, students, policymakers and other stakeholders. In the future, educational institutions can use the Learning Education Check System in order to analyze the progress of their students, employees and education in general and base their future decisions driven on real-time data.

In the upcoming time we search for reliable and valid methods to analyze the behavior of individual actors. The stadium in which the social sciences nowadays is, asks for the connection of instruments to analyze big data relevant in educational settings. However, in the future we should offer a more integrated form of research, where one instrument conjoin all possibilities (the ideal LECS)

To use the educational big data in a self-regulating and sustainable system, it is important to reflect on the reliability and validity issues that automatically arises in a 'new' research topic like educational big data. The infrastructure for the monitoring and evaluation of learning education uses a large amount of data, which differs in nature. The reliability of the data depends on the contents of the data and the analysis. The reliability of for example the questionnaires, panel research and interviews can be measured with traditional methods for reliability, as Cohen's Kappa and the Cronbach's Alpha. But since big social data is quite a new phenomenon, these general reliability measurements should be discussed. Corbin (1986, p. 102) indicates that the analysis of raw data (such as big social data) is a complicated process in which the data is reduced and converted to concepts, which are designed for representing categories, when following the grounded theory. One way to do this is to develop a code book in which a clear explanation of labelling the relevant big social data makes it possible to carry on the inter-subjectivity of the inter-rater reliability of a test. In this case, the Cohen's Kappa used in qualitative research can be used in the big social data analysis too. For quantitative data within the big social data, the Cronbach's Alpha can be used.

To improve the validity of the qualitative and quantitative parts of the data in the infrastructure, it is important that the researcher indicates how the structuring of the data collected has been established (Boeije, 2008). Potter and Levine-Donnerstein (1999) describe two steps to ensure the validity. The first step is to develop a coding scheme that coders use during the content analysis, wherein the scheme is considered as reliable if it is closely related to the theory. A second step is to determine the agreement between coders, a high degree of similarity produces coding valid data.

The next steps are to develop and implement the LECS in practical situations. Besides that it is necessary to do more scientific research to the opportunities and pitfalls of educational big data for practical and scientific use.

References

- Boeije, H. (2008). *Analyseren in kwalitatief onderzoek. Denken en doen*. Den Haag: Boom onderwijs.
- Brown, J. S., & Adler, R. P. (2008). Open education, the long tail, and learning 2.0. *Educause review*, Vol. 43, No. 1, pp. 16-20.
- Cambria, E., Rajagopal, D., Olsher, D., & Das, D. (2014). Big social data analysis. *Big Data Computing*, pp. 401-414
- Carroll, M. W. (2006). Creative Commons and the New Intermediaries. *Mich. St. L. Rev.*, 45.
- Carver, C. S., & Scheier, M. F. (1982). Control theory: A useful conceptual framework for personality-social, clinical, and health psychology. *Psychological bulletin*, Vol. 92, No. 1, pp. 111.
- Carver, C. S., & Scheier, M. F. (1999). *Control theory: A useful conceptual framework for personality-social, clinical, and health psychology*. In R. F. Baumeister, et al. (Eds.), *The self in social psychology*. Philadelphia: Psychology Press/ Taylor & Francis.
- Corbin, J. (1986). *Coding, writing memos and diagramming*. In Chenitz, W.C., & Swanson J.M., (eds). *From practice to grounded theory: qualitative research in nursing*. Menlo Park, Addison-Wesley.
- Constantinides, E., & Fountain, S. J. (2008). Web 2.0: Conceptual foundations and marketing issues. *Journal of Direct, Data and Digital Marketing Practice*, Vol. 9, No. 3, pp. 231-244.

- Deming, W.E. (1950). *Elementary Principles of the Statistical Control of Quality*, JUSE
- Doole I. & Lowe R. 2004. *International Marketing Strategy*. Thomson Learning
- Franklin, T., & Van Harmelen, M. (2007). Web 2.0 for content for learning and teaching in higher education. *JISC* www.jisc.ac.uk/media/documents/programmes/digitalrepositories/web2-contentlearningand-teaching.pdf.
- García-Peñalvo, F. J., Conde, M. Á., Alier, M., & Casany, M. J. (2011). Opening learning management systems to personal learning environments. *Journal of Universal Computer Science*, 17(9), 1222-1240.
- Haythornthwaite, C., De Laat, M., Dawson, S., & Suthers, D. (2013, January). Introduction to Learning Analytics and Networked Learning Minitrack. In *System Sciences (HICSS), 2013 46th Hawaii International Conference on*. IEEE. pp. 3077-3077
- Hesselbein, F., Goldsmith, M., & Beckhard, R. (1997). *The Organization of the Future*. San Fransisco: Jossey-Bass Publishers
- Kop, R. (2011). The challenges to connectivist learning on open online networks: Learning experiences during a massive open online course. *The International Review of Research in Open and Distance Learning*, 12(3), pp.19-38.
- Marsh, J. A., Pane, J. F., & Hamilton, L. S. (2006). Making sense of data-driven decision making in education.
- Mazza, R., & Dimitrova, V. (2004, May). Visualising student tracking data to support instructors in web-based distance education. In *Proceedings of the 13th international World Wide Web conference on Alternate track papers & posters*. ACM. pp. 154-161
- McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big Data. *The management revolution*. *Harvard Bus Rev*, Vol. 90, No. 10, pp. 61-67.
- Pang, J., Jie, L., & Xu, F. (2014). Study on the Group Cooperative Innovation Based on WEB2. 0. *Journal of Software*, Vol. 9, No. 3, pp. 613-620.
- Pedler, M., Burgoyne, J. & Boydell, T. (1991). *The learning company*. London: McGraw-Hill.
- Potter, W.J. & Levine-Donnerstein, D.(1999). Rethinking validity and reliability in content analysis. *Journal of Applied Communication Research*, 27(3), pp. 258-184. doi: 10.1080/00909889909365539
- Pullen, J. M., Brunton, R., Brutzman, D., Drake, D., Hieb, M., Morse, K. L., & Tolk, A. (2004). Using Web services to integrate heterogeneous simulations in a grid environment. In *Computational Science-ICCS 2004* (pp. 835-847). Springer Berlin Heidelberg.
- Raban, D.R., & Rafaeli, S. (2006). Investigating ownership and the willingness to share information online. *Computers in Human Behavior*, 23(2007), pp. 2367-2382.
- Senge, P.M. (1990). *The fifth discipline: The art and practices of the learning organization*. Doubleday Currence: New York.
- Slater, S.F., & Narver, J.C. (1995). Market orientation and the learning organization. *The journal of Marketing*, Vol. 59, pp. 63-74.
- Teece, D. J. (2010). Business models, business strategy and innovation. *Long range planning*, 43(2), pp. 172-194.
- West, D. M. (2012). Big data for education: Data mining, data analytics, and web dashboards. *Governance Studies at Brookings, September*. <http://www.brookings.edu/~media/research/files/papers/2012/9/04%20education%20technology%20west/04%20education%20technology%20west.pdf>.
- Wiener, N. (1948). *Cybernetics: Control and communication in the animal and the machine* (pp. 194). New York: Wiley.
- Zikopoulos, P., & Eaton, C. (2011). *Understanding big data: Analytics for enterprise class hadoop and streaming data*. McGraw-Hill Osborne Media.

Appendix 1: Big social data for educational purposes

	Educational big data Analysis method	Educational big data sensor	Analytics	Analysis	Data collection Software	Educational big data analysis software	Data visualization software
1	Simple Educational big data Analytics	Message Counter	i.e. Popularity Reputation Influence	Overview of the number/percentages of messages that were sent.	General educational big data monitoring tools (i.e. Coosto, Radian6, Social Mention)	i.e. Nvivo AtlasTI Dedoose MAXQDA	GenStat JMP Matlab NCSS
2		Messenger Counter	i.e. Activity	Overview of the number/percentage of messagers that has sent a relevant message.	General educational big data monitoring tools (i.e. Coosto, Radian6, Social Mention)	Coding Analysis Toolkit HyperRESEARCH	SAS SPSS Stata
3		Messenger Profiler	i.e. Audience knowledge	Overview of the messagers' profiles.	General educational big data monitoring tools (i.e. Coosto, Radian6, Social Mention)		Note: Plenty of options
4		Reaction Profiler	i.e. Popularity Reputation Influence	Analysis of the rate of posts and comments	General educational big data monitoring tools (i.e. Coosto, Radian6, Social Mention)		
		Social search counter	i.e. Popularity Reputation Influence	Analysis of the rate of searches	Google Trends (simple social search analytics)		
5	Educational big data Content Analytics	Structured Theme Profiler	i.e. Theme development Genre development	Timelines	Scraper (Google-plugin) OutWit Hub Screen-Scraper	i.e. Nvivo AtlasTI Dedoose MAXQDA	Nvivo AtlasTI SPSS (and other statistical programs)

6		Unstructured Theme Profiler	i.e. Opinion mining, Sentiments	Clusters of messages	General educational big data monitoring tools (i.e. Coosto, Radian6, Social Mention) Scraper (Google-plugin) OutWit Hub Screen-Scraper	Coding Analysis Toolkit HyperRESEARCH	GenStat JMP Matlab NCSS SAS SPSS Stata Note: Plenty of options
7		Theme Persona profiler	i.e. Opinion mining	Content analysis	Scraper (Google-plugin) OutWit Hub Screen-Scraper		
8		Theme Coding Technique	i.e. Opinion mining	Labeling/coding	Scraper (Google-plugin) OutWit Hub Screen-Scraper		
9	Educational big data Network Analytics	Timeline Network development	i.e. Network development (over time)	Network analysis	NodeXL	Messages and networks from Facebook and Twitter	<p>Gephi: Viewing and analyzing network graphs</p> <p>NodeXL (Excel): Viewing and analyzing network graphs</p> <p>Visone: Analysis and visualization of social networks</p>
10		Sub-Network profiler	i.e. Single Facebook Fanpage Multiple Facebook Fanpage	Cluster analysis (based on time stamp)	ScraperWiki	Messages and followers from Twitter	GenStat JMP Matlab NCSS SAS SPSS Stata Note: Plenty of options
11		(Social) Relationships	i.e. Inbetweenness Centrality Density Cohesion	Cluster analysis (based on posts and responses)	General educational big data monitoring tools (i.e. Coosto, Radian6, Social Mention)	Messages from all educational big data	

			Subgroup Identification Brokerage Co-likes & Co-comments					
12	Cross analytics	media	Cross-Media Content profiler	i.e. Cross references	Content analysis of media	OutWit Hub Screen-Scraper	ATLAS TI NVIVO	ATLAS TI NVIVO

*The software applications in the matrix are just a small sample