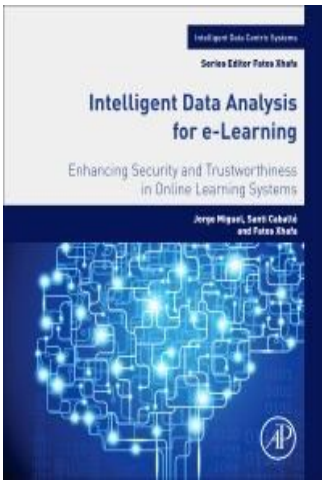
BOOK REVIEW

INTELLIGENT DATA ANALYSIS FOR E-LEARNING
Enhancing Security and Trustworthiness in Online Learning Systems
Written by Jorge Miguel, Santi Caballe, and Fatos Xhafa

Hilal Seda YILDIZ AYBEK
Independent Researcher
Learning & Development Specialist
Eskisehir, Turkey

ISBN	**978-0-12-804535-0**
Publication Date	**2017**
Publication Formats	**Hardcover and e-Book**
Publisher	**Elsevier**



With the rapid development of Internet technologies, various paradigms of learning can be adapted to e-learning environments. One of these paradigms, Computer-Supported Collaborative Learning (CSCL), can be presented to learners through web-based systems such as LMS while incorporating peer-to-peer (P2P) learning, measurement, and evaluation strategies. In this book titled *Intelligent Data Analysis for e-Learning Enhancing Security and Trustworthiness in Online Learning Systems*, various strategies and applications are presented to ensure trustworthiness in e-learning environments, especially where the CSCL paradigm is adopted. A comprehensive literature review on student security, privacy, and trustworthiness has been presented in a very detailed and comprehensive way. This allowed readers to conceptually prepare for detailed applications in the later parts of the book and case studies at the Universitat Oberta de Catalunya. In addition to the applications that are presented in detail, the approaches and techniques such as Learning Analytics, Educational Data Mining, distributed computing, and massive data processing are shared through detailed applications of how to adapt to the measurement and evaluation applications offered in online learning environments in the context of trustworthiness.

In the first chapter, *Introduction*, concerns about the web security for e-learning environments used by many educational institutions have been mentioned. The most important of these concerns are security threats that may arise in measurement and evaluation tools or modules in web-based learning environments such as LMS. The author mentions the lack of preventive work on web security. Especially in online collaborative learning environments, emphasized that there is a need for planned and definitive solutions for trustworthiness.

In the second part, *e-Learning Security*, recent studies on privacy and security in e-Learning environments have been addressed. In particular, it emphasized the need to establish common policy-based privacy and security management standards on how student privacy can be maintained and protected by existing systems. Recent studies have emphasized the necessity of adopting risk-oriented approaches to confidentiality in e-learning and recording student activities in this context. Determination of the motivations of invasion of privacy, the development of guarantor-mediated reputation, and context-based identity schemes which do not require the storage of the student's specific data are also included in recent studies. To prevent a threat, it is emphasized that the category in which the threat belongs must be identified. In this context, security threats are classified into three categories: the sources dimension, the techniques dimension, and the results dimension. Public Key Infrastructure (PKI) based solutions that manage certificates and key pairs are also handled in this section. The most common security attacks in LMS are categorized as *virus attack, denial of service, sniffing network, denial of involvement, web spoofing, man-in-the-middle attack, credential spoofing, session hijacking, timestamp, and brute force attack*. Security strategies for a paradigm based Collaborative Learning, Mobile Learning, and Massive Open Online Courses are also presented in this section for securing LMSs that are used to provide interaction, cooperation, and coordination between learner-learner, learners-instructor, and learner-content.

In the third chapter, *Trustworthiness for secure collaborative learning*, Trustworthiness is defined and summarized under three headings, basic trust, general trust, and situational trust. The factors affecting the trust score were examined by presenting the factors and rules developed to determine, measure, and model the level of trustworthiness. It is emphasized that an LMS needs to include useful learning resources and reliable peer services to be trustworthy. Since trustworthiness strategies are often P2P-based, it has been stated that these strategies have been facilitated in adapting to structured environments through the CSCL paradigm in e-learning systems. In the second part, the author refers to the development of strategies based on student data, and in this section, he provided hints on how these e-Learning data can be interpreted and analyzed. In this part, detailed cross-sections are presented from the present study on knowledge management in particular. Several strategies have been proposed for the analysis of detailed and large data on learning behaviors in the LMS log files. One of these is the use of the Apache Hadoop distributed system, where MapReduce model can be run. The author, who summarized the techniques and methods of Educational Data Mining, then mentioned data visualization and briefly introduced EDM tools such as GISMO, LOCO-analyst, PSLC DataShop and Meerkat-ED, and SNA visualization tools such as Cytoscape and Gephi. Another sub-title, Trustworthiness-based CSCL, focuses on functional security models that can be used for CSCL. CSCL has been mentioned before regarding the trustworthiness of P2P in common terms; In this part, the security issue in e-Assessment is discussed concerning availability, integrity, identification and authentication, confidentiality and access control, and non-repudiation features. Finally, the e-Exam case study presented via LMS was introduced and it was discussed how the exam could be improved regarding security. As a result of this case study, a remarkable result has been reached that security issues cannot be completely solved only by technology.

In the fourth chapter, *Trustworthiness modeling and methodology to secure peer-to-peer e-Assessment*, the stages of the trustworthiness model and the notation and terminology used are introduced. Quantitative data collection tools used to feed trustworthiness purposes and model are listed as ratings, questionnaires, P2P evaluation questionnaires, student reports, and LMS usage indicators. The functions used in the summarized normalization process are explained in detail. After normalization, trustworthiness levels are modeled, and the Pearson correlation analysis processes used to determine whether the variables included in the model relate to each other are described in detail. Data analysis, which has an important place in trustworthiness and security methodology, is covered in the context of Educational Data Mining and Data Visualization; Prediction, clustering, outlier detection, relationship mining, social network analysis, and process

mining are suitable solutions for trustworthiness evaluation and prediction. Another subheading in this section is the creation of student profiles in the e-Assessment. The author has emphasized the need to develop a Collective Intelligence application for the creation of student profiles. At this stage, the use of manageable student profiles in LMS has been suggested. These profiles have features that allow students to see the profiles of their peers, access, and edit student profiles; of course, administrators and system administrators, and users can edit their profiles. In this e-Assessment system employed by Collective Intelligence, the continuous assessment approach has been adopted, and at this stage, it has planned to maintain continuity using questionnaires, forums, and P2P surveys. The last part of this section is the adaptation of the Trustful Student Profile approach to a MOOC platform. This model, called MOOC-SIA, is suitable for all types of courses and user profiles, as well as real-time user tracking for student verification, allowing large data to be processed with different data mining techniques. However, it has been stated that this approach requires high computational power, therefore, is costly.

In the fifth chapter, *Massive data processing for effective trustworthiness modeling*, explanations are given on the parallel programming models used for processing large data. It has been stated that the use of parallel models rather than sequential is more advantageous and cost-effective in the problem of "high computational power required for processing large amounts of data" mentioned in the previous section. Although the trustworthiness-based approach for e-Learning environments, which is discussed in detail in Chapter 3, is effective for Internet Security, it is emphasized that it can be costly for real-time processing of very large data. Therefore, a parallel processing model has been developed and tested. The Hadoop MapReduce implementation, which is implemented for use in the P2P student data analysis process, is described in detail.

In the sixth chapter, *Trustworthiness evaluation and prediction*, the three cases of the Universitat Oberta de Catalunya referred to earlier in this book were evaluated and predicted regarding trustworthiness. In the first study called CSCL-course-1, the continuous assessment model was developed and studied in four stages with 12 students. The second study, called P2P-course, was conducted in a seven-level course with 57 students. The final study, called CSCL-course-2, has the same design as the first one, but with different strategies developed against situations that arise when the first study is performed. Factors for carrying out of trustworthiness evaluation of these applications were determined, and statistical analysis and normalization procedures were performed on these data and reported. For trustworthiness prediction, the Neural Network approach, which captures non-linear relations, is used and the results of the analysis are reported.

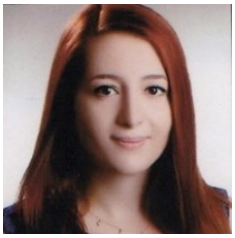
In the seventh chapter, *Trustworthiness in action: Data collection, processing, and visualization methods for real online courses,* the processes dealt with in the previous sections and handled separately have been adapted to real online courses. The research instruments and the technological tools made use of, are explained in detail. Detailed information about the virtual environment in which Apache Hadoop is run is shared. Also, the development process of MapReduce applications, the technical features of the installed system, the code fragments, and commands used in these processes are explained in detail. Performance analysis for processing model evaluation was performed and reported through data visualization. For P2P e-Assessment, various data visualization processes were performed with NodeXL software.

In the eighth chapter, *Conclusions and future research work*, the strategies and applications developed in the previous sections on Trustworthiness in security in computer-supported collaborative learning, and the methods and techniques used are summarized. Anomalous user assessment, P2P visualization tools, Trustworthiness prediction, and massive data processing are discussed. The information related to the experimentation and validation procedures related to these subjects is briefly shared. Finally, the results obtained from the developed applications are briefly discussed.

Intelligent Data Analysis for e-Learning Enhancing Security and Trustworthiness Online Learning Systems **is a well-organized book that guides practitioners in the e-Learning field to adopt new security and trustworthiness approaches to their systems. Explanations of the code and scripts used in the book, especially in the application parts, made the book quite understandable and prevented the reader from having a question mark in their minds.**

In the context of e-Learning systems, different strategies are presented to improve the security, privacy, and trustworthiness of the systems and to improve the systems. At the same time, readers are advised on how to apply them. It can be described as an important source of information in e-Learning, Information Security, and Intelligent Data Analysis, since the book is written not only for theories but also for practice, and contains detailed information on how to configure and use various tools.

BIODATA and CONTACT ADDRESSES of AUTHOR



Hilal Seda YILDIZ AYBEK is a Ph.D. student in Distance Education Department of Anadolu University Social Sciences Institute. YILDIZ AYBEK completed her master's education in the same department in 2016. She took her BS degree in Computer and Instructional Technology Education Department in 2014. YILDIZ AYBEK currently works as a freelancer Learning & Development Specialist. YILDIZ AYBEK's research areas include open and distance learning technologies, instructional design, user experience, learning support services, intelligent and adaptive learning systems, learning management systems, artificial neural networks, and educational data analysis.

Hilal Seda YILDIZ AYBEK
Phone: +905445041382
E-mail: hilalsedayildiz@gmail.com

REFERENCES

Miguel, J., Caballe, S., & Xhafa, F. (Eds.). (2017). ***Intelligent Data Analysis for e-Learning Intelligent Data Analysis for e-Learning Enhancing Security and Learning Systems*****. Boston: Academic Press.**