WHO RUNS THE WORLD: DATA

Editors
Sevinç GÜLSEÇEN, Sushil SHARMA, Emre AKADAL



İSTANBUL UNIVERSITY PRESS

WHO RUNS THE WORLD: DATA

EDITORS

Prof. Dr. Sevinç GÜLSEÇEN
Prof. Dr. Sushil SHARMA
Dr. Emre AKADAL

İSTANBUL UNIVERSITY PRESS

Published by
Istanbul University Press
Istanbul University Central Campus
IUPress Office, 34452 Beyazıt/Fatih
Istanbul - Turkey

www.iupress.istanbul.edu.tr

WHO RUNS THE WORLD: DATA
Editors: Sevinç Gülseçen, Sushil Sharma, Emre Akadal

e-ISBN: 978-605-07-0743-4

DOI: 10.26650/B/ET06.2020.011

Istanbul University Publication No: 5273
Informatics Department Publication No: 5

Published Online in December, 2020

It is recommended that a reference to the DOI is included when citing this work.

This work is published online under the terms of Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)
https://creativecommons.org/licenses/by-nc/4.0/

This work is copyrighted. Except for the Creative Commons version published online, the legal exceptions and the terms of the applicable license agreements shall be taken into account.

CONTENTS

CHAPTER 1
MATERIAL INFORMATION CARRIERS: HISTORICAL DEVELOPMENT
Róbert JÁGER 1

CHAPTER 2
ASTRONOMICAL DATA
Hulusi GÜLSEÇEN, Hasan H. ESENOĞLU 13

CHAPTER 3
DATA STORAGE IN THE DECENTRALIZED WORLD:
BLOCKCHAIN AND DERIVATIVES
Enis KARAARSLAN, Enis KONACAKLI 37

CHAPTER 4
DATA IN THE CONTEXT OF INDUSTRY 4.0
Fatma Önay KOÇOĞLU, Denizhan DEMİRKOL 71

CHAPTER 5
BIG DATA GOVERNANCE
Malgorzata PANKOWSKA 93

CHAPTER 6
A CORE PROBLEM WITH HUMAN DATA PROCESSING: EPISTEMIC
CIRCULARITY IN ACTION
Mehmet Selim DERİNDERE 107

CHAPTER 7
DATA PRE-PROCESSING IN TEXT MINING
Tuğçe AKSOY, Serra ÇELİK, Sevinç GÜLSEÇEN 123

CHAPTER 8
BIG DATA IN EDUCATION: A CASE STUDY ON PREDICTING E-LEARNING
READINESS OF LEARNERS WITH DATA MINING TECHNIQUES
Zeki ÖZEN, Elif KARTAL, İlkim Ecem EMRE 145

CHAPTER 9
THE VALUE OF DATA FOR IMPROVING EFFECTIVENESS OF
CAMPUS COURSES: THE CASE OF HYBRID MOOCS
Oğuz AK, Selim YAZICI, Sevinç GÜLSEÇEN 165

CONTENTS

CHAPTER 10
INTELLIGENT TUTORING OF LEARNERS IN E-LEARNING SYSTEMS AND MASSIVE OPEN ONLINE COURSES (MOOC)
Yacine LAFIFI, Asma BOUDRIA, Atef LAFIFI, Moadh CHERAITIA 177

CHAPTER 11
SMART HOUSE: DATA GATHERING AND ANALYSIS
Natalija LEPKOVA 193

CHAPTER 12
PRIVACY FOR ENTERPRISES IN THE DATA AGE
Bilgin METİN, Enes YILMAZ, Erdi ŞEKERCİLER 209

CHAPTER 13
DATA COLLECTION APPROACHES FOR ARTIFICIAL INTELLIGENCE APPLICATIONS IN HEALTHCARE
Murat GEZER, Çiğdem SELÇUKCAN EROL 227

CHAPTER 14
THE TECHNOLOGICAL TRANSFORMATION PROCESS FROM ELECTRONIC INTELLIGENCE TO CYBER INTELLIGENCE
Ahmet Naci ÜNAL 239

CHAPTER 15
AUTOMATIC MEASUREMENT OF THE MORPHOLOGICAL CHARACTERISTICS OF HONEYBEES WITH A COMPUTATIONAL PROGRAM
Zlatin ZLATEV, Veselina NEDEVA, Ivanka ZHELYAZKOVA 261

ABOUT THE CONTRIBUTORS

Dr. Sevinç Gülseçen, is a Professor and Head of the Department of Informatics at Istanbul University and also Director of Computer Applications and Research Center, Istanbul, Turkey. She has her bachelor degree in math and astronomy from Faculty of Natural Sciences and her doctoral degree in artificial neural networks from Faculty of Business, Istanbul University.

Dr. Gulsecen has her primary teaching and research interests in system analysis and design, constructivist learning, e-learning, computer-mediated communications, e-government, community and social informatics, and knowledge management. She has published her research in more than 50 papers in several national and international journals, conference proceedings and edited books such as International Journal of E-adoption, IEEE Technology and Society Magazine, Educational Technology and Society, Euroasia Journal of Mathematics, Science and Technology Education, International Review of Research in Open and Distance Learning, Education and Information Technologies, Technologies for Enhancing Pedagogy, Engagement and Empowerment in Education: Creating Learning-Friendly Environment.

Dr. Gülseçen has been member of Istanbul University International Academic Relations Board, Istanbul University Institute of Science Academic Board and Turkish Informatics Society. Dr. Gulsecen has served on the editorial board of several refereed journals and has reviewed manuscripts and conference proceedings. She has served as a moderator and session chair at national and international conferences. She was a Turkey delegate to USA, England, Italy, Germany, Poland, Slovakia, Lithuania, Bulgaria and China for her professional meetings and work. She has also worked with business professionals for her consulting and research.

Dr. Gülseçen is a co-founder and co-chair of the international conference named "FutureLearning: Innovations in Learning for the Future".

Dr. Sushil Sharma, is currently an associate dean and a professor of computer information systems (CIS) in the Miller College of Business at Ball State University. Dr. Sharma has co-authored/edited/co-edited twelve (12) books and published over ninety (90) refereed research papers in the most reputed national and international information systems' journals. In addition, he has published forty-five (45) refereed chapters in various books and has presented and published over 160 papers in various national and international conferences. Dr. Sharma's research has appeared in several highly ranked journals in the MIS field, including Decision Support Systems, Communications of the Association for Information Systems, European Journal of Information Systems, Information Systems Frontiers, Journal of Information Privacy & Security (JIPS), Electronic Commerce Research Journal, and Information Management and Computer Security.

His primary research interests are in computer information systems security, e-Learning, e-Government, computer-mediated communications, human computer interaction (HCI) and community and social informatics. He has taught several graduate and undergraduate courses on a variety of subjects including database management, ERP systems (SAP), electronic commerce, computer and network security, management information systems, systems analysis and design, distributed data processing systems, computer networking, knowledge management and other information systems related subjects. Since 2007, Dr. Sharma has served as the editor-in-chief/co-editor-in-chief of the International Journal of E-Adoption. He has been a guest editor, associate editor, reviewer and member of the editorial boards for several national and international journals in the area of MIS.

ABOUT THE CONTRIBUTORS

Dr. Emre Akadal, is a researcher at the Department of Informatics at Istanbul University. He has a bachelor's degree in Physics from the Faculty of Science and a master and doctoral degree in Informatics from the Institute of Graduate Studies in Science at Istanbul University.
His primary research interests are in evolutionary algorithms, database design, internet-based programming, and human-computer interaction. Dr. Akadal has research papers over ten in related national and international journals. He has served as the co-editor of the International Journal of Acta INFOLOGICA, since 2017. Also, he has been the manager of Istanbul University Human-Computer Interaction Laboratory. Dr. Akadal has been taken part in the team of several projects about information management.

Doc. JUDr. PhDr. Róbert Jáger, PhD. PhD. studied law, philosophy and cognitive sciences. Pedagogically, he focuses on teaching the legal history and legal philosophy. Scientifically, he focuses mainly on the development of the oldest law in the territory of Central Europe. Author uses linguistic methods to reconstruct the law at a time when there was no written record of legal norms. He currently points to the need for critical thinking in interpreting information and text.

Assoc. Prof. Hulusi Gülseçen, graduated from Astronomy and Space Science Department of Faculty of Science of Ege University in 1980. He received his Master of Science degree in 1983 and his PhD degree in 1989 from Institute of Graduate Sciences of Istanbul University with the theses on "The Commensurability Problem in Asteroids" and "The Effects of Corriolis Force in Sunspots" respectively. Dr. Gulseçen is currently working at Astronomy and Space Sciences Department of Faculty of Science of Istanbul University. His primary researh and teaching interest is in the topics of double stars, data analysis and astroinformatics. His research has appeared in journals such as Publications of the *Astronomical Society of Australia, New Astronomy, Education and Information Technologies*. He is lecturing both in Astronomy and Space Sciences Department and Informatics Department of Istanbul University respectively. In 2011, Dr. Gülseçen was visiting schollar at Astronomy Department of Ball State University in Muncie, Indiana (USA).

Assoc. Prof. Hasan Hüseyin Esenoğlu, (PhD) was born in 1963 in Eskisehir (Turkey). After his graduation from Istanbul University, Faculty of Science, Department of Astronomy and Space Sciences in 1982, he started his MSc in Physics Education by attending the English preparatory class in Marmara University. He completed his master's degree in Astrophysics in 1990 at the same department of Istanbul University with the thesis titled "Observational Properties of Two Cataclysmic Binary Stars". After completing his PhD at the same university, he started to work as guest astronomer at the Asiago Astrophysical Observatory of Padova University (Italy) under the scholarship of the Italian Ministry of Foreign Affairs. By preparing a thesis titled "Spectral and Photometric Analysis of Novae: A New Classification Method", he graduated from Stellar Astrophysics PhD program in 1996. Dr. Esenoğlu also visited King Saud University Science Faculty at Astronomy Unit of Physics Department (Riyat, Saudi Arabia) as a lecturer for three years, and the National Observatory of The Scientific and Technological Research Council of Turkey (TUBITAK) (TUG, Antalya, Turkey) as the Chief Researcher for five years. Dr. Esenoğlu primary research interest is on classical novae, cataclysmic variable stars and meteorites. He actively continue follow-up observations with ground based telescopes for alert sources of GAIA

and SRG (in the near future) satellites. His research has appeared in well known Astronomy journals (SCI) and he also published papers on popular science.

Enis Karaarslan, is currently Assistant Professor at the Mugla Sitki Kocman University (MSKU) Computer Engineering Department. He established NetSecLab for network/security education and research. He has been studying the potentials of the blockchain in BcRG (Blockchain Research Group) since 2017. He is one of the founders of the Artificial Intelligence (AI) discipline in the MSKU. His research areas are computer networks, cybersecurity, blockchain, data science, disaster management, and digital twin. He has over 50 papers to his name. He likes biking, hiking, traveling and reading.

Enis Konacaklı, has been managing CIS and IT teams, systems, services, from acquisition through delivery on through their lifecycle sustainment for 18 years. He is currently the Director of CIS of Eskişehir Air Base. He has been studying the potential benefits of the blockchain technology for National security and Military implementations as a member of BcRG (Blockchain Research Group) since 2017. His research areas are computer networks, cybersecurity, blockchain.

Fatma Önay Koçoğlu, (PhD) graduated from Istanbul University (IU) Faculty of Science, Department of Mathematics and Sakarya University Faculty of Engineering, Department of Industrial Engineering. With her dissertations on Data Mining, she received her master's degree in 2012 and doctorate in 2017 from IU Institute of Graduate Studies in Sciences, Department of Informatics. In 2011, she was a graduate student at Université de Technologie de Compiégne (France), Computer Engineering Department for 6 months. She is continuing her second doctorate in IU-Cerrahpaşa Institute of Graduate Studies, Industrial Engineering Program. She has been working as a Research Assistant at IU Informatics Department since 2010. Her research areas are Artificial Intelligence, Data Mining, Machine Learning, Decision Support Systems, Facility Location and Allocation Problems, Mathematical Modeling and Optimization. She has many scientific studies published in national and international fields. She worked as a researcher in projects supported by The Scientific and Technological Research Council of Turkey (TÜBİTAK) and the other Higher Education Institutions. She also worked as an instructor in the scientific training programs organized by Higher Education Institutions and Non-Governmental Organizations. She speaks English fluently and basic French.

Denizhan Demirkol, received his bachelor's degree in Information Systems and Technologies from Yeditepe University. He completed his master's degree in Yeditepe University, Department of Management Information Systems with honors degree. Demirkol continues his doctorate education at Istanbul University, Institute of Science and Engineering, Department of Informatics. He has been working as a research assistant at Aydın Adnan Menderes University Department of Management Information Systems since 2016. He has been working as a research assistant at Istanbul University Informatics Department due to his doctorate education. He gave lectures in the field of "Human Computer Interaction" at "SMK University of Applied Social Science" within the scope of Erasmus teaching mobility. During his PhD education, he started his studies in Data Science and continues to study Human –Computer Interaction, Artificial Intelligence, Machine Learning and Data Mining.

ABOUT THE CONTRIBUTORS

Malgorzata Pankowska, is Professor of Social Science and Chair of the Department of Informatics at the University of Economics in Katowice, Poland. She received the qualification in econometrics and statistics from the University of Economics in Katowice in 1981, the Ph.D. degree in 1988 and the Doctor Habilitatus degree in 2009, both from the University of Economics in Katowice. She was visiting professor at ISLA Braganca in Portugal, Trier University in Germany, ICHEC in Brussels, Belgium, VGTU in Vilnius, Lithuania, Istanbul University in Turkey, Ionian University in Corfu, Greece, Universidad de Ibague, Ibague-Tolima, Colombia, and Lapland University in Kemi-Tornio, Finland. She is in the Board of Information System Audit and Control Association (ISACA) Katowice Chapter. List of publications and more information on website: http://web.ae.katowice.pl/pank/

Mehmet S. Derindere, works with executives and teams to develop effective problem solving practices using knowledge. He has worked in several private organizations as trainer, business analyst, project manager, auditor and manager. Currently interested in cognitive and organizational traps that ensnare individuals as well as teams, prevent effectiveness and reduce quality of life.

Tugce Aksoy, graduated with a B.A. in Translation and Interpreting from Istanbul University in June 2018. She is currently studying in M.Sci program in Informatics at Istanbul University.
She completed her scientific preparation program and attended various courses such as statistics and programming during that period.
She works as a Vendor Manager in a translation company and continues to do academic work. Tugce has also worked as a translator for several years.
Tugce's research topics are NLP (Natural Language Processing), text mining, and statistics.

Serra Çelik, took her Phd in Quantitative Methods Department at School of Business in Istanbul University. She currently works in Department of Informatics in Istanbul University. Serra's research topics are multivariate data analysis, statistical learning and data mining. She gives lectures master level courses as Data Analysis, Quantitative Methods in Business. Serra recently works on Digital Transformation and Natural Language Processing. She has a web site about her academic life (serra.ist).

Zeki Özen, is a research assistant at Istanbul University, Informatics Department. He graduated from Sakarya University Computer Programming Department in 2002 and Istanbul University, Faculty of Science, Department of Physics in 2008. He received his MSc in 2012. In 2013, Dr. Ozen was at Ball State University, Miller College of Business, Information Systems and Operations Management (IN, USA) as a visiting scholar for 2 months. He received his Ph.D. degree in 2016 from Istanbul University, Informatics Department with his thesis titled *"Developing a security system for authentication based on keystroke dynamics using artificial neural networks"*. Research areas of Dr. Ozen are artificial intelligence, machine learning, data mining, biometric authentication systems, relational database systems, and source code comparison/similarity.

ABOUT THE CONTRIBUTORS

Elif Kartal, started her undergraduate education at Istanbul University (IU), Faculty of Science, Department of Mathematics in 2004. In 2008, she started her master's degree in IU Informatics Department and was accepted to the same department as a Research Assistant. She completed her master's degree in the IU Informatics Department with her thesis titled *"Software project cost estimation with artificial neural networks"* in 2011, and her doctorate with her thesis titled *"Machine learning techniques based on classification and a study on cardiac risk assessment"* in 2015. In 2013, Dr. Kartal was at Ball State University, Miller College of Business, Information Systems and Operations Management (IN, USA) as a visiting scholar for 2 months. From August 2019 to March 2020, Dr. Kartal was a visiting scholar at the University of Miami, College of Arts and Sciences, Department of Computer Science (FL, USA), and worked on computational neuroscience, deep learning, and computer vision with Assoc. Prof. Odelia Schwartz (Ph.D.). This postdoctoral research supported by TUBITAK-BIDEB 2219-International Postdoctoral Research Scholarship Program. Machine learning, data mining, and artificial intelligence are the main research areas of Dr. Kartal.

İlkim Ecem Emre, is a research assistant at Marmara University, Management Information Systems Department. She graduated from the same department in 2015. She has started the graduate program at Istanbul University Informatics Department in 2015 and graduated in 2017, where she wrote a thesis titled *"Analysis of effects of acute rheumatic fever in childhood on heart disease with data mining".* She is currently a Ph.D. student at İstanbul University Informatics Department. Her research areas are data mining, machine learning, and medical informatics.

Oğuz Ak, Ph.D., Instructor at Bogazici University, During his education period he get various courses like educational technology, pedagogy, information technology and informatics. He has a B.S. in Computer Education and Educational Technologies & M.A. in Management Information Technologies at Bogazici University. And Ph.D. in Informatics at Istanbul University. He worked as a distance education coordinator in a private university for a year. He has been working as an academician at Bogazici University more than 10 years. He is currently making researches about distance education, educational technologies and digital games.

Selim Yazıcı, Ph.D., Professor at Istanbul University, Selim YAZICI is Professor of Management and Organization at Department of Business Administration at Istanbul University, Faculty of Political Sciences. He received his BSc in Mechanical Engineering from Yildiz Technical University and his MSc and PhD in Business Administration from Istanbul University. He is also Adjunct Professor at Marmara University, Banking and Insurance Institute, Department of Insurance (2009-2016) and Ozyegin University, Graduate School of Business, Financial Engineering and Risk Management Program (2018-…). He is lecturing on E-Learning, Entrepreneurship, Financial Technologies, Digital Insurance, Project Management, *Business Continuity Management*, Management, Organizational Behavior and International Business both at undergraduate and graduate levels. His research interests are *E-Learning*, Digital Transformation, Disruptive Technologies, Entrepreneurship, Startups, Financial Technologies (FinTech), Insurance Technologies (InsurTech), Business Continuity Management, Team Building, Leadership and Experiential Learning. He is the author of four books published in Turkish: Business Continuity

Management (2013), E-Learning (2004), E-Insurance (2002), and *Learning Organizations* (2001). He is the co-founder of *FinTech Istanbul*, a platform for supporting financial technology (FinTech) startups and building a strong ecosystem in Turkey. He has also designed a "FinTech 101 Training Program" for entrepreneurs, corporates, and investors which is the first program of its kind in Turkey. Selim YAZICI is a member of Turkish Quality Association (KalDer), Academy of International Business (AIB) and The Chamber of Mechanical Engineers (TMMOB).

Yacine Lafifi, is currently working as a Full Professor at the Computer Science Department of Guelma University, Algeria. Since November 2018, he is the Vice-Rector of Development, Foresight and Orientation at Guelma University. Also, he is a senior researcher at LabSTIC laboratory (Guelma University, Algeria) where he is the leader of the "Web Technology and Intelligent Systems" team. He works in e-learning research field since 1997. He received his PhD in computer science from the University of Annaba (Algeria) in 2007. He has several published papers in conferences and journals. Furthermore, he is an editorial board member of many international journals (International Journal of Web-Based Learning and Teaching Technologies, Journal of Theoretical and Applied, SAGE Open journal, Electronic Journal of e-Learning, Acta INFOLOGICA, Frontiers of Contemporary Education, and International Journal of Information and Communication Technology Education). Currently, he works on e-tutoring environments, e-Learning, CSCL, recommender systems, Artificial intelligence in education, Intelligent agents, e-technology, MOOC and human tutoring systems.

Asma Boudria, is currently a PhD student in computer science at LabSTIC Laboratory (Guelma University, Algeria). She has worked for three years for implementing massive open online courses at LabSTIC laboratory. She is preparing a PhD thesis about the learning and evaluation in MOOC. Her current research concerns include e-learning systems, MOOC and human tutoring systems, personalization in MOOC and social learning in MOOC.

Atef Lafifi, obtained his master's degree in Computer Science in June 2017 from the University of Guelma (Algeria), specialty: Computer Systems. He has a good knowledge of website development in general and e-learning systems in particular. He masters several programming languages such as JAVA and PhP. He worked on tutoring in MOOCs.

Moadh Cheraitia, obtained his master's degree in Computer Science in June 2017 from the University of Guelma (Algeria), specialty: Computer systems. He masters several programming languages such as JAVA and PhP. He worked on tutoring in MOOCs.

Natalija Lepkova, doctor of technological sciences, Associate Professor at the Department of Construction Management and Real Estate, Civil Engineering Faculty, Vilnius Gediminas Technical University, Lithuania. She has 17 years of experience in teaching at university, specializing in facilities management, quality management systems, BIM in facility management. She published 2 books, 6 chapters in books and a number of journal papers in the mentioned fields. Participated in international projects, such as INTEREG Project LONGLIFE, also COST. She is a member of the editorial board of the following journals: International Women Online

ABOUT THE CONTRIBUTORS

Journal of Distance Education; Turkish Online Journal of Distance Education; Baltic Journal of Real Estate Economics and Construction Management, BIM Arabia. She is a member of: GYODER, IFMA Czech Republic Chapter, IFMA, Lithuanian construction engineers union, National Association of Distance Education (NADE).

Bilgin Metin, firstly received the B.Sc. degree in Electronics and Communication Engineering from Istanbul Technical University, Istanbul, Turkey. Later, he worked in the leading companies in the private sector, and he handled the design and installation of computer networks and cybersecurity projects. He received M.Sc. and Ph.D. degrees in Electrical and Electronics Engineering from Bogazici University, Istanbul. During M.Sc. and Ph.D. studies, he worked as a consultant in the private sector for designing, implementing, and supporting data communications and cybersecurity projects such as the International Terminal of Antalya Airport in Turkey.
In 2007, He started to work as an assistant professor for Management Information Systems Dept. at Bogazici University. He received the title of Associate Professor in 2014. His research interests include cybersecurity, IT Governance, data privacy, and telecommunication circuit design. He published more than 100 papers in international journals and conferences. Assoc. Prof. Bilgin Metin has been on the organization committee of many national and international conferences.
He has been also the head of the Bogazici University MIS Cybersecurity Center Project since 2017. He is carrying out national and international projects, also giving training and consultancy to various national and international institutions.

Enes Yılmaz, was graduated from Bogazici University, Management Information Systems Department in 2017. Currently, they continue their career in the information technology area.

Erdi Şekerciler, was graduated from Bogazici University, Management Information Systems Department in 2017. Currently, they continue their career in the information technology area.

Murat Gezer, has received PhD degree from Istanbul University, Natural Sciences Institute Electrics and Electronics Engineering. He is a working in Istanbul University Informatics Department. His teaching and research interests focus on Image Processing, Signal Processing, Computer Vision, datamining and data science programming. Furthermore detailed CV of Dr. Gezer, which has many national and international publications, book chapters and projects, can be accessed from https://www.researchgate.net/profile/Murat_Gezer

Dr. Çiğdem Selçukcan Erol, received her PhD degree from İstanbul University, Institute of Science in 2010. Dr. Erol is working in Istanbul University Informatics Department since 2003. She received the title of associate professor in Management Information Systems in 2018. Her research interests in bioinformatics, data mining, machine learning and artificial intelligence in health. Furthermore detailed CV of Dr. Erol, which has many national and international publications, books and projects.

ABOUT THE CONTRIBUTORS

Ahmet Naci Ünal, graduated from Gazi University. Then, he got his master's and PhD degrees at the Defense Technologies Department of the Institute of Science and Technology at Istanbul University. Between the years 1991-2012, he worked as an officer lecturer in the Turkish Air Force. He retired in 2012 at his own request. Ahmet Naci ÜNAL started to work as an assistant professor at Bahçeşehir University in 2013. He is also the Director of the Cyber Security Implementation and Research Center at Bahçeşehir University and the Coordinator of the Cyber Security Graduate Program in the Graduate School of Natural and Applied Sciences at Bahçeşehir University. He is specialized in electronic-based defense technologies, decision support systems and cyber security concepts. He has various books, articles and symposium papers published in these fields.

Zlatin Zlatev, is an Associate Professor at the Department of Technics and technologies, Trakia University, Bulgaria. He holds a Bachelor's degree in Electrical Engineering at the Trakia University and a master's and a doctorate in Process automation at the Ruse University, Bulgaria. Research interests: Computer-Aided Design, Informatics, Image processing, and analysis. Assoc. Prof.Zlatev is author of more than 100 publications and more than 95 citations. List of publications and more information on the website: https://www.researchgate.net/profile/Zlatin_Zlatev

Veselina Nedeva, PhD, is an Associate Professor in Informatics and Computer Science at Trakia University – Stara Zagora, Faculty of Technics and Technologies – Yambol, Bulgaria. She graduated Master's degree in Informatics from the University of Economics – Varna, and Ph.D. in Application of Computing in Economics at the same University. A dedicated research fellow with thirty years of experience in computer science, program languages, databases, information systems, computational data, e-learning, and distance learning. Extensively published in science journals and scientific conferences and author of more than 105 scientific publications and 130 citations.

Ivanka Zhelyazkova, is Professor of Apiculture and Head of sub-depatrment Beekeeping and Sericulture of the department of Animal husbandry-Nonruminants and other animals at the Faculty of Agriculture in Trakia University, Stara Zagora, Bulgaria. She received the PhD degree in 1999 and the DSc degree in 2013, both from the Trakia University in Stara Zagora. Main scientific interests: Biology of the honeybee and technologies for raising bee colonies; Oportunities for using bees and bee products as bioindicators for pollution of the environment. Prof. Zhelyazkova is author of more than 100 publication and more than 230 citations.

PREFACE
Who Runs the World: DATA

In ancient times, possessing a land and gold was the most important asset in the world. In the 21st century, "Data" is becoming the most important asset. As data activities continue to increase in speed, scale and variety, data Science and Data Analytics is becoming the new phenomenon of the 21st century. In fact, Data is considered to be the gold of the 21st Century and "data" is changing the face of our world. The organizations are becoming bigger and bigger and need a large amount of data to process for creating a business insight or intelligence. As more and more companies worldwide start offering products and services online, companies need to not only process information quickly but also to get insight about the needs, expectations, transactional behavior and responses of the customers. This mandates the companies to deal with "Data" effectively. Nowadays, "Data" is in various forms such as; texts, documents, online books, music, videos on number of different platforms such as; social media, transactional websites and many other internet based online forums.

The "Data" explosion is creating newer exciting opportunities for companies and individuals at the same time, it is also creating concerns and challenges. This edited book presents the research work of several researchers who are working in the data science related areas. The book comprises 15 chapters. A brief abstract of each chapter is provided below:

Material Information Carriers: Historical Development
Róbert Jáger, Matej Bel University, Slovakia

In this study, we will briefly try to describe what material information carriers were like in each period of the development of human society, what the advantages or disadvantages of these carriers were, and how society changed with the change of the material information carriers themselves. In conclusion, we will highlight an interesting fact of current development: the digitization of material information carriers, the separation of the information itself from its material carrier, and the risk that it will face in the future. In the first part of this study we describe Material information carriers in prehistoric times and in antiquity. Specifically, we focus on the development and use of material information carriers in Ancient Egypt, Mesopotamia and compare them with the development in European communities, which reached a similar level of civilization in the later period. In the second part, we pay attention

to the issue of using material information carriers in the Middle Ages. We show how the material carriers changed in the given period. In the third part of this study, we pay attention to the issue of using material information carriers in modern times. In particular, we point out the risks of the current state of information storage in information systems. The study primarily uses methods of description, analysis, synthesis, comparison, abstraction, and generalization.

Astronomical Data

Hulusi Gülseçen, İstanbul University, Turkey
Hasan H. Esenoğlu, İstanbul University, Turkey

Space telescopes have increased the quality of data collection for today's astronomy. In parallel to this, obtaining high quality data with high technology and good resolution focal plane detectors in accordance with the developments in material science in the ground-based observations has been achieved. With the new generation of ground based and space observations, global campaigns also brought continuity in data acquisition and increased performance. Finally, the fact that theoretical outputs can be made to allow in today's technology, for example, the detection of gravitational waves in the universe and these add new ones to the existing data. In addition, there has been a significant increase in data archiving, reduction and processing together with the number and variety of data collection tools. Astronomers have been able to overcome the facilitation in these processes in their own way: manpower waste has been reduced with autonomous telescopes, the data has been transformed into informatics (astroinformatics) with pipelines, the workload has been reduced to large masses by establishing a virtual observatory, and finally smart applications have been opened with the provided big data and new open areas have been reached with a future such as data mining. In this way, there has been progress in solving many astronomical events in the universe. This chapter is orginized in two subsections. In first, we are discussing how to solve problems in astronomy by using big data. In the second, we mention about big data sources in astronomy. The importance of data in astronomy, sources of data, big data in regards to the discovery of universe and analyzing data are the topics discussed in these subsections.

Data Storage in the Decentralized World: Blockchain and Derivatives

Enis Karaarslan, Mugla Sitki Kocman University, Turkey.
Enis Konacaklı, Eskisehir Technical University, Turkey.

We have entered an era where the importance of decentralized solutions has become more obvious. Blockchain technology and its derivatives are distributed ledger technologies that keep the registry of data between peers of a network. This ledger is secured within a successive over looping cryptographic chain. The accomplishment of the Bitcoin cryptocurrency proved that blockchain technology and its derivatives could be used to eliminate intermediaries and provide security for cyberspace. However, there are some challenges in the implementation of blockchain technology. This chapter first explains the concept of blockchain technology and the data that we can store therein. The main advantage of blockchain is the security services that it provides. This section continues by describing these services.. The challenges of blockchain; blockchain anomalies, energy consumption, speed, scalability, interoperability, privacy and cryptology in the age of quantum computing are described. Selected solutions for these challenges are given. Remarkable derivatives of blockchain, which use different solutions (directed acyclic graph, distributed hash table, gossip consensus protocol) to solve some of these challenges are described. Then the data storage in blockchain and evolving data solutions are explained. The comparison of decentralized solutions with the lcentralized database systems is given. A multi-platform interoperable scalable architecture (MPISA) is proposed. In the conclusion we include the evolution assumptions of data storage in a decentralized world.

Data in the Context of Industry 4.0

Fatma Önay KOÇOĞLU, İstanbul University, Turkey
Denizhan DEMİRKOL, Aydın Adnan Menderes University, Turkey

Today, every sector, not least industry, has been affected by the development of technology. With the breakthrough development of technology, Industry 4.0 has emerged with the concept of big data. Data is the most important element in the process of creating information. This study aims to deal with the subject of Industry 4.0 which has attracted great interest in the global field in the context of big data. Studies concerning Industry 4.0 and related data are examined in our study through a systematic literature review. Web of Science database and “industry 4.0 and data” keywords were used for our article search. A preliminary evaluation was performed for 20 articles meeting the objective of this study which were selected for

detailed examination. When the studies on Industry 4.0 and data are analyzed, we can determine that studies with big data, digitalization, internet of things, digital twin, cyber-physical systems, smart factories and cloud computing are prominent. Moreover, when the countries where the articles were published were analyzed, it was found that China was the most cited and studied country in this field. It is believed that the results of this examination will enlighten people working in this field and direct future studies.

Big Data Governance

Malgorzata Pankowska, University of Economics in Katowice, Poland

Information processing in a traditional way focuses on relatively stable structured data, repeatable processes as well as on operations in Business Intelligence systems. However, nowadays more and more popular, big data, defined as huge volumes of data available in varying degrees of complexity, generated at different velocities, and varying degrees of ambiguity, cannot be processed using traditional methods and technologies. Some people argue that suitable IT (Information Technology) infrastructure for big data processing is not yet widely developed nor implemented to discuss the big data architecture implementation benefits, risks, and opportunities. Nevertheless, this paper is to present the big data governance issues. Particularly, within the proposed theme, the author discusses the big data system architecture and development strategy. The last part of the paper includes a proposal of a big data architecture model as well as a design of balanced scorecard objectives and measures specification to support the big data governance at public services business organizations. As usual, there are two main research methods, i.e., literature review and the analysis of case studies. The first provides an overview of the existing knowledge and the second permits for contextualization of the proposed models. Beyond that, the paper includes definitions of the key concepts and enables to extend the knowledge base in the research area.

A Core Problem with Human Data Processing: Epistemic Circularity in Action

Mehmet Selim Derindere, İstanbul University, Turkey

Managers are expected to solve critical problems of our society in an efficient manner and in ways so that the problems remain solved. In order to accomplish this, the managers are provided with vast amounts of resources including mountains of data and a wide variety of problem-solving methods available. On the other hand, the effectiveness of social and organizational problem solving is far from satisfactory and this lack of effectiveness is

ubiquitous. One reason of this ineffectiveness we claim has to do with how the human mind works. The inherent capabilities and limitations of human mind coupled with social-cognitive skills lead to sub-par problem solving. An especially counterproductive problem solving approach used by managers is setting and attempting to solve problems using erroneous cognitive skills that not only fails to include relevant data but also uses the existing data in a counterproductive manner. The very data processing skills of managers make problem solving a dead end for the actors involved at great cost to them and to the society.

This chapter looks at a core human data processing problem that renders the available data and techniques ineffective. Epistemic Circularity disregards all the disconfirming or threatening data and fails to include it in the problem solution. Epistemic Circularity thus renders the revelant data useless in developing effective solutions. Easy knowledge, a product of epistemic circularity, leads to ineffective problem solving which in many cases result in exacerbated problems and counterproductive consequences.

Data Pre-processing in Text Mining

Tuğçe Aksoy, İstanbul University, Turkey
Serra Çelik, İstanbul University, Turkey
Sevinç Gülseçen, İstanbul University, Turkey

The fact that any kind of user has the ability to generate data with great ease at any time causes an increase in the importance of data mining. Considering the reality that the vast majority of the available data is composed of unstructured data and that the data in the text type is outnumbering, it proves the increasing interest in text mining and the abundance of studies in this field. However, in order to be able to examine an unstructured data type like text, which is quite different from machine language, it is necessary to make this data more structured and make the machine work. At this point, the data pre-processing step, which covers a large part of the entire text mining process, is of great importance. In this chapter, it is aimed to explain the text pre-processing phase on a basic level by supporting this using visuals. In doing so, it is primarily planned to focus on text mining and to explain in detail the characteristics of the data processed. In this context, it is aimed to explain the data pre-processing steps followed in order to overcome these difficulties by examining the difficulties created by the data in question. As a result, this chapter is a descriptive review of the data pre-processing phase in text mining, which covers some of the studies previously conducted on this subject.

Big Data in Education: A Case Study on Predicting E-learning Readiness of Learners with Data Mining Techniques

Zeki Özen, İstanbul University, Turkey
Elif Kartal, İstanbul University, Turkey
İlkim Ecem Emre, Marmara University, Turkey

Since the term "personalized learning" became popular, smart features have begun to be integrated into the e-learning environment. Data mining and machine learning algorithms are used to analyze big data stored in an e-learning system to make predictions to improve course quality or learners' performance. From the learners' perspective, it might now be considered possible for everybody to benefit from e-learning by considering their personal interests or their own specific development plan as long as the course contents are available in the system. In addition, in an e-learning environment there is no limitation on the time and place where a course can be attended and a program completed. However, it is just not that simple. Today not the only, but by far the most important, requirement is still the readiness of the learners to study in an e-learning system. The aim of this chapter is to predict the e-learning readiness of learners using data mining techniques and to provide feedback for institute managers and admin personnel of e-learning systems which are intended to be used in an institution. According to the results of this study, the highest accuracy value (0.831) is obtained with C4.5 Decision Tree Algorithm. While students, who agree and strongly agree with the statement "My studying/research area is appropriate for e-learning" are classified as ready to attend an e-learning course, students who disagree with the same statement are classified as not ready to attend an e-learning course. Students who strongly disagree with the statements "My studying/research area is appropriate for e-learning" and "E-learning is better than face to face learning", are also classified as not ready to attend an e-learning course.

The Value of Data for Improving Effectiveness of Campus Courses: The Case of Hybrid MOOCs

Oğuz Ak, Boğaziçi University, Turkey
Selim Yazıcı, İstanbul University, Turkey
Sevinç Gülseçen, İstanbul University, Turkey

In recent years with the advances in technology, learners started to learn various concepts in informal learning environments apart from the official traditional learning programs.

We describe such learning environments as part of the Personal Learning Environment (PLE) approach. One great resource for these environments is using Massive Open Online Course (MOOC). Learners can learn any subject by enrolling in MOOCs easily and develop themselves by reaching their personal learning goals. But in such an informal learning environment, it would be hard to manage the learning process. Learners need some ability to manage this process that is called "self-regulation". There seem to be some problems in both fully face to face learning (like difficulties in following courses), and fully online MOOCs (like lack of interaction). So, a midway approach is a hybrid MOOC that is a combination of both methods. Literature and author experiences indicated that this method would make learning more effective. However, there is a need for improving the method with proper data management. We provide a list of data collection methods in hybrid MOOCs and explain how this data helped us to improve the learning process. In the PLE approach, students need data to shape their learning process, similarly instructors need to obtain data with various strategies and reshape the course structure by using this data. We think that in education, data usage is somehow limited, but it is required for making it more efficient.

Intelligent Tutoring of Learners in E-learning systems and Massive Open Online Courses (MOOC)

Yacine Lafifi, University 8 May 1945, Guelma
Asma Boudria, University 8 May 1945, Guelma
Atef Lafifi, University 8 May 1945, Guelma
Moadh Cheraitia, University 8 May 1945, Guelma

In the last few years, many terms related to learning environments have emerged. Each one of these terms is distinguished by a set of criteria such as the target audience, the duration of learning, the type and nature of the educational content, the manner of dissemination of knowledge, etc. Unfortunately, lack of support for the learners seems to be a serious frequently faced problem in these environments, which requires special attention. Among proposed solutions, tutoring seems to be a convenient candidate for this problem. Tutoring involves offering assistance to learners that are in need for help. Regardless the nature of this assistance (pedagogical, social, etc.), it can be delivered in many forms: advice, guidance or even recommendation. And while tutoring has been applied for decades in traditional e-learning environments, its application in new systems such as Massive Open Online Courses (MOOC) is still under study. In fact, a considerable number of studies driven on MOOCs had reported the problem of learners' dropout. Several reasons can be listed as

causes of such a problem. Among these reasons, we can find learners' isolation as well as learners' loss of motivation. This same problem has been reported by researchers working in the field of Computer-based Environments for Human Learning.
In this article, we propose a new vision on how to apply an intelligent tutoring process in human learning systems in general and in MOOCs in particular. This new vision is based on the behaviors and skills of learners. This activity can take many forms and can be carried out by different types of actors (teachers, learners, etc.).

Smart House: Data Gathering and Analysis
Natalija Lepkova, Vilnius Gediminas Technical University, Lithuania

In modern society, the concept of "smart house" is increasingly being heard. At present, it is generally acceptable that a smart house has efficient building management, local management and business management systems. A smart house increases the business value of the environment created by the adaptability and flexibility provided by the location and the communication systems. There are many opinions on how we should understand the concept of a smart house. Some people believe this is a modern home audience, others think it's a fully designed home cable system. There are some who guess that this reflects modern telecommunication systems, etc. Everything that has been mentioned really reflects only part of the "smart home" possibilities. Smart House introduces a modern, robust automated system that allows to integrate all of the main operating subsystems such as: energy supply, supply of gas and water, lighting system, heating systems, microclimate systems, other remote controls. The smart houses are often pointed as one of the main constituents of smarter living environments. The chapter provides the smart house definition, criteria defining smart building, smart house technology explanation, examples of smart houses in different countries, smart house data model, building progress and analysis of smart building automation and control systems (applying SWOT analysis method).

Privacy for Enterprises in Data Age
Bilgin Metin, Bogazici University, Turkey
Enes Yılmaz, Bogazici University, Turkey
Erdi Şekerciler, Bogazici University, Turkey

The world we live in is now becoming increasingly virtual. We all interact with this new age which we can describe as the digital age. We shop online, we communicate with people via

social media, we are informed at any time through the devices that are in our hands about goings-on, whether we like it or not, we have become a part of this globalized and digitalized world. Data can be described as the structure of the digitalized world. In each interaction between us and the tools which we use, we create data or we cause data transferring or we can be a small part of a large data collection because of our presence in a platform on the internet. Certainly, this close relationship can reveal our private life in some situations. Most of the time, we are exposed to situations where our private information is collected, used, and processed without our permission. Sometimes we cannot even notice the violation of one of the most fundamental rights and freedoms we can define as privacy. This literature survey study is based on the fundamentals of information security, and it seeks answers to these questions: Why does our personal information need protection? What kind of information should be protected? What is the situation regarding the data privacy in Turkish and world law? What kind of laws have been passed upon the privacy of tax from past to today? What are the perspectives, opinions on protection of personal data in Turkey and Europe? What is the importance of data privacy for the business sectors? We also believe that this study will raise awareness on this matter.

Data Collection Approaches for Artificial Intelligence Applications in Healthcare

Murat Gezer, İstanbul University, Turkey
Çiğdem Selçukcan Erol, İstanbul University, Turkey

As in all other fields, research in the field of artificial intelligence is rapidly continuing in the field of health. As a result of this research, the importance of data comes to the fore. In this study, which includes data collection approaches in the field of health, we aim to emphasize the importance of data in this field and to contribute to the more conscious handling of the data to be used in artificial intelligence applications at every stage. For this purpose, the definition of data and how to distinguish information and knowledge are mentioned. The characteristics of data and data collection methods are also mentioned, and an attempt is made to emphasize the importance of health data collection in artificial intelligence research. As a result of this study, we believe that all personnel working in data-related departments and the health field, where the moment is vital, must receive training on collecting, storing, sharing data, and data security in particular. In our study we emphasize that especially the people who produce and consume data must have the awareness and morality for every step of data collection and handling, and that this issue should be prioritized in the field of health.

Technological Transformation Process from Electronic Intelligence to Cyber Intelligence
Ahmet Naci Ünal, Bahcesehir University, Turkey

Throughout the whole of human history, concepts such as defense, security, safety and intelligence have been very important for human beings on a personal level, and for human communities in general. The sociological transformation that came as a result of these processes played a key role in the development of science and technology. Thanks to the developments in electronic science during the 20th century, systems using electromagnetic energy have come to the fore.

This development process which started with systems such as telephones, radios and radars has eventually been used in many different areas such as air defense systems, guided missiles, early warning receivers, communication systems, and computers. For this reason, the control and active use of what can be called the electromagnetic spectrum in short, has been an important factor in all kinds of activities. By the 21st century, almost all of the systems used in this process began to operate in a cyberspace environment and became software controlled. The concept of the target intelligence needed in this transformation process has changed dimensions and shifted from electronic intelligence to cyber intelligence.

This study will focus on the transformation of the electronic intelligence process, which is an indispensable element of the 20th century, into the concept of cyber intelligence in the 21st century.

Automatic Measurement of The Morphological Characteristics of Honeybees by A Computational Program
Zlatin Zlatev, Trakia University, Bulgaria
Veselina Nedeva, Trakia University, Bulgaria
Ivanka Zhelyazkova, Trakia University, Bulgaria

The use of Big data related to the breeding of honey bees, when administered and processed effectively, will encourage the development of knowledge-based beekeeping, create new markets and business opportunities and further encourage the development of this industry. There have been attempts to fully automate the process of measuring the morphological characteristics of bees (at this stage there are conversions for Measuring wings), but this process for other parts are still completed manually. A survey was made of the possibilities

to automate the process of measuring the morphological characteristics in honeybees and the proposed algorithm and program to implement it. Color characteristics of parts of the bee body - tergite and proboscis, through which they can be separated from the background of the image, are analyzed and measured. Distances are determined between the values of the colour components of the object and background. From statistical analysis, it is found that S and V colour components of the HSV colour model are appropriate for the separation of an object from the background . Algorithms and a program in Matlab environment for separating tergite and proboscis from the background of the image and definition of their main sizes are developed. From the analysis of the results, it is found that the major influence on the accuracy of the measurement is of the bee in the image.

We are deeply indebted to several colleagues and students who have helped to complete this book. We owe a special debt of gratitude to several reviewers who provide valuable suggestions to authors for improving their manuscripts. We also acknowledge helpful advice from Istanbul University Press. I am sure readers will like the contributions of various researchers presented in this book and in last but not the least, we would like to thank the sections' authors, the management and staff of Istanbul University for their support and contributions in the emergence of this work.

CHAPTER 9

THE VALUE OF DATA FOR IMPROVING EFFECTIVENESS OF CAMPUS COURSES: THE CASE OF HYBRID MOOCS

Oğuz AK*, Selim YAZICI, Sevinç GÜLSEÇEN*****

*Ph.D., Boğaziçi University, Faculty of Education, Computer Education and Educational Technology, İstanbul, Turkey
E-mail: oguz.ak@boun.edu.tr

**Prof, İstanbul University, Faculty of Political Sciences, Department of Business Administration, İstanbul, Turkey
E-mail: selim@istanbul.edu.tr

***Prof, İstanbul University, Informatics, Informatics, İstanbul, Turkey
E-mail: gulsecen@istanbul.edu.tr

DOI: 10.26650/B/ET06.2020.011.09

Abstract

In recent years with the advances in technology, learners started to learn various concepts in informal learning environments apart from the official traditional learning programs. We describe such learning environments as part of the Personal Learning Environment (PLE) approach. One great resource for these environments is using Massive Open Online Course (MOOC). Learners can learn any subject by enrolling in MOOCs easily and develop themselves by reaching their personal learning goals. But in such an informal learning environment, it would be hard to manage the learning process. Learners need some ability to manage this process that is called "self-regulation". There are some problems in both fully face to face learning (like difficulties in following courses), and fully online MOOCs (like lack of interaction). So, a midway approach is a hybrid MOOC that is a combination of both methods. Literature and author experiences indicated that this method would make learning more effective. However, there is a need for improving the method with proper data management. We provide a list of data collection methods in hybrid MOOCs and explain how this data helped us to improve the learning process. In the PLE approach, students need data to shape their learning process, similarly instructors need to obtain data with various strategies and reshape the course structure by using this data. We think that in education, data usage is somehow limited, but it is required for making it more efficient.

Keywords: Hybrid massive open online courses, Massive open online course, Educational data analysis, Data for evaluation, Data for quality of education

1. Introduction

In recent years, online learning degrees have been created in addition to traditional classroom learning programs. Many people have been learning with the help of computers and the internet for decades, and they have gotten degrees. The Personal Learning Environment (PLE) approach -apart from the institutional degrees- is another way of learning. In this approach, learners set their learning goals and reach the learning content that they need and learn by themselves (Initiative, 2009). The last generation seems to use this method as well (Hollands & Tirthali, 2014).

Of course, there are some challenges in using this method. One of the main challenges is that students need to have self-regulation skills. Self-regulation is -in brief- a skill for managing individual learning processes (Zimmerman, 2000). Learners need to determine what they are required to learn and how to manage that process. Today, learners that have this skill can reach almost all knowledge within structured mediums like massively open online courses (MOOCs). MOOCs are online courses that can usually be accessed for free, without any prerequisites and are offered by credible universities around the world. Currently, there are really huge amount of MOOC resources that can easily be reached and millions of students are taking them (Liyanagunawardena, Adams, & Williams, 2013; Shah, 2015). Moreover, it is also known that the current generation can learn differently (Prensky, 2001). In addition, they can manage their learning process easily by using various tools like the Learning Management Systems (LMS).

Although some students can successfully manage their personal learning environments, some cannot. Similarly, although some students perform well in traditional learning courses, some do not. To help poorly achieving students in both environments a hybrid approach should be used. In particular, some online materials can be implemented into learning courses and both type of students could benefit from it. In this chapter we will demonstrate the use of hybrid Massive Open Online Courses (h-MOOCs) (Perez-Sanagustin, Hilliger, Alario-Hoyos, Kloos, & Rayyan, 2017).

The aim of this chapter is to clarify a personal learning environment method and how it could support traditional learning in the h-MOOC form. One important requirement for managing this process and maximizing its effectiveness is the use of good data management. It is because, data usage will enlighten learners and the course providers in such informal learning environments. Different types of data could be collected in h-MOOC applications with various methods like observations, interviews with students, logs from the tools etc. So,

we will discuss how to realize the inclusion of MOOCs into classroom learning and improve the process with the use of data.

2. An Alternative Learning Approach: Personal Learning Environments

2.1. What is Personal Learning Environment (PLE) Approach?

The personal learning environment approach consists of tools, services, and communities that create a learning environment with the aim of reaching students' personal learning goals (Initiative, 2009). So, to learn anything, participants do not have to follow a degree path, but can learn by themselves through this approach. Actually, this method existed before the computer and internet era, but the current tools and services make it a more feasible way of learning. And it seems people are using this method.

Since the invention of computers and the internet, learning management systems (LMS) have been using. With LMS, instructors can manage learning periods and create valuable data to better follow students' development. Personal Learning Environment (PLE) applications may be similar to LMS, but actually they are different. LMS is primarily designed for course management but PLE's main aim is to manage the learning period (Ebner & Taraghi, 2010; Initiative, 2009). It is important to differentiate the PLE approach and tools as well. PLE tools are software designed for using as part of the PLE approach.

To better understand this method, we need to discuss the place where learning occurs. Traditionally schools were addressed as the primary learning place. However, a school is not the only area, there is significant learning happening outside schools (Humanante-Ramos, García-Peñalvo, & Conde-González, 2015). Students can learn from other sources like books, media, friends, and family. To support this fact, Banks et al. (2007) stated that students only get about 19% of all learning gains in the classroom within their first 12 school years. A significant amount of learning occurs outside. In the PLE approach, our aim is to better manage the outside learning. Students need to have the ability to manage their own learning periods by using quality learning resources. In outside learning, descriptive data is needed for delivering these good information resources to students. In this way, learners could better decide their learning requirements and find the related resources. For example, when anyone buying a product from the internet he should follow some steps: First, he needs to decide what is required and then analyzes the products to see if the quality and specifications fit the requirements. In online sales, most stores try to provide data that best describes the product, they try to get feedback from users to describe how well the product functions. Manufacturers may use this data to produce better products next time. The PLE approach is similar to this

marketing example. Learning is a requirement and a person needs to have the ability to manage their learning needs and find resources that best fit their requirements. So, using good data in the PLE approach is critical.

Currently there are different physical and digital learning resources, but MOOCs seem to be the one that is well structured (but still need more development) and easy to access. So, in a PLE approach learners could use various learning resources including MOOCs. In this chapter we will try to better understand MOOC resources and in particular their integration with face to face courses using data.

2.2. Self-Regulation in Personal Learning Environments Approach

Self-regulation is an ability that includes thoughts, emotions, and actions to help a learner reach their personal goals (Zimmerman, 2000). Recently the term was mostly related with online environments (e.g. Kuo et al., 2014). It is because, students need to have more autonomy and responsibility skills in an online learning environment (Artino, 2008). Normally if students learn from instructors face to face, it is easier to communicate and ask questions, but in an online environment they have limited interaction with the content provider. It seems, it is one of the most difficult part of this approach, but good self-regulation skills could solve this problem.

Moore (1993) defined the "transactional distance" concept as the psychological and communicative distance between learner and instructor, and it is clearly high in online learning environments. However, in such environments like MOOCs this distance can be closed by a leaners' autonomy skill (Shearer, Gregg, Joo, & Graham, 2014). So self-regulation is a critical skill. To get this skill, students need to learn how to study by themselves from an early age. It seems once students have this skill they can manage learning by themselves for their whole life. We think 'learning how to dress' as an example: First, a child gets instructions of how to dress from an early age. Parents give instructions on how to put on a piece of clothing step by step, and never allow the child to try their own way. With this learning approach, every time the child faces a new type of dress, they would have difficulty and require parent support. In this example a parent's help is similar to an instructors' presentation in the classroom. Students need to learn their own way. If they don't find methods to learn by themselves, they would require an instructor's help in every learning process and this will limit learning.

3. Massive Open Online Courses (MOOCs)

3.1. A Good Resource for Learning

If we consider the profiles of current personal learners, we are faced with two major types of learners. The first type is the adult learners that are trying to improve their job skills (Castano-Munoz, Kreijns, Kalz, & Punie, 2017) and the second type is the students that are trying to complete their school learning. Both of them require information which they could get from various tools including the internet. Of course, there are other learner profiles as well, for example parents who need to learn "how to care for their child" or a person who would like to learn "how to cook" etc.

There are lots of learning opportunities today. Historically, learning has evolved with the advances in technology. Especially TV, radio, computers, and the internet have affected learning processes (Liyanagunawardena et al., 2013). After MIT's open Courseware project in 2001 a new era of "open learning" was started. It was because previously reaching good courses or course materials was limited as they were part of some learning degrees. In 2008 the first massive open online course was created (Liyanagunawardena et al., 2013) and since then there has been a massive open online course movement around the world.

Today, more than five thousand MOOCs are created, and more than 60 million learners are already registered for them. So, they are being actively used. Currently, there are two types of MOOCs; xMOOCs and cMOOCs. xMOOCs usually consist of video lectures and evaluated assignments (Siemens, 2013) and today most of the MOOCs in the USA are in this form (Daniel,2012). On the other hand, cMOOCs use a connectivist approach, which includes social interaction and online tools to create knowledge networks (Siemens, 2013). Any learner can reach a MOOC from any part of the world with internet access usually for free. It is a good opportunity because the MOOC content is provided by top universities in the world.

Hollands and Tirthali (2014) asked the reasons for providing MOOCs to MOOC administrators. They listed a set of reasons that includes; increasing the reach to learning, improving brand, economic reasons, innovation, and research about education. Interestingly, they did not talk much about the use of data or improving the quality of learning by using it. This would be the one missing point most educationalists have about this method of learning. Learning is not a process of content providing and presentation, and administrators need to acquire better use of the data inside the courses.

3.2. MOOC in Personal Learning Environment Approach

We stated that the PLE approach needed tools, services, and communities. As tools, online learning environments provide some software that includes LMS functionality. In addition, there are lots of services and communities that support learners like google docs, google drive, forums, discussions networks etc. Any person can use them to achieve their personal learning goals, but it may be hard for some learners to manage lots of different components. Especially for more traditional learners, a MOOC could be a better alternative because they are designed as a whole.

Any learner can take a MOOC to reach their own personal learning goal. Some platforms even offer MOOC degrees that are not like traditional learning programs. They are a kind of learning path recommended for some main learning targets. Learners can follow these paths and improve their ability in a specific field. The difference between these paths and traditional ones is that they are not mandatory, learners do not have to follow a strict schedule, and they have a certain prerequisite to access the learning content including entrance exams. As a result, MOOC is a good learning resource for a personal learning environment approach.

4. Hybrid Approach with the Use of Data

4.1. Hybrid Massive Open Online Courses to Improve Classroom Learning

Blended learning is a method for using both face to face and online learning together. Means et al. (2009) stated that blended learning usually results in better outcomes than solely online or solely face to face learning. So, although personal learning is usually outside official learning programs, they could be implemented into them, because in reality learners already try to use online resources to support their school learning.

There are different types of blended learning approaches. Kloos et al. (2015) stated the ways as: local digital prelude, flipping classroom canned digital teaching with remote tutoring, face to face and connected teaching, live teaching with remote tutoring, and face to face teaching with remote tutoring. Similarly, Perez-Sanagustin et al. (2017) showed four types of MOOC integration in the classroom as: MOOC as a service (MOOC can be recommended and not related to curriculum), MOOC as an added value (a MOOC that is related with curriculum is recommended), MOOC as a replacement (face to face course is replaced by a MOOC), and MOOC as an operator (face to face course is operated by a MOOC).

There are some h-MOOC examples in the literature. Konstan et al. (2015) applied a "MOOC as an added value" type of implementation. The application resulted in good learning gains and in general better student perception. Similarly, Bruff et al.(2013)'s "MOOC as an operator" application study showed that students similarly liked this application. In other studies, researchers reported usually more positive effects than the negative ones (Robinson, 2016; Swinnerton, Morris, Hotchkiss, & Pickering, 2017). They all showed that students in the current generation would benefit from this application.

Moreover, in one study, students used some common PLE devices and more than 90% of them stated that they would continue to use the platform after graduation (Tsui & Sabetzadeh, 2014). It seemed that PLE tools supported lifelong learning.

4.2. Data Collection in Hybrid MOOC Applications

In MOOCs, there were too many learners, so it was hard to give personal feedback and evaluation needed to automatize (Daradoumis, Bassi, Xhafa, & Caballé, 2013). Current automatic evaluation systems have limited interaction and seem not good enough. Similarly, peer review -which is a valuable evaluation type--is limited again because the learners were not professionals (Daradoumis et al., 2013). However, we think that automatic evaluation systems could evaluate learners, the problem is not related to its nature, but it is related to the limitations of current technologies.

Liyanagunawardena et al. (2013) analyzed 45 studies and found data collection techniques such as email interviews, focus groups, Moodle (an LMS) logs, discussion forums, blogs, and observations. By using a combination of techniques, MOOCs can be better evaluated. But in the case of hybrid MOOCs, data collection is somehow different. Actually, it is hard to reach MOOC usage data, but with alternative data collection methods students' data could be collected. We think it is better to evaluate what was learned than what was reached. For example, to understand student gains from a MOOC chapter, a very basic assignment about the related MOOC content could be given to students. To do so, short answers or writing some comments for discussion forum entries would be good.

While in MOOCs some quantitative data like questionnaires are collected, in hybrid MOOCs there is an opportunity to use qualitative measurements as well, because there is more interaction with student and instructor in face to face sessions. For example, the instructor can ask for feedback about the MOOC content and make observations about students understanding.

5. A Case Study: Data Analysis from a Hybrid Massive Open Online Course

In our case, an introductory database management course was given to 3 different groups of higher education students with a hybrid MOOC approach for three consecutive years. In each group there were about 30 students (there were 92 students in three years, 41% female and 59% male) who made good scores in the national university entrance exam in Turkey. Participants were in their 3rd year at university (average age is about 23). Participants followed half of the course from a MOOC while they followed the other half face to face. The MOOC course was an introductory database course that was given by Stanford University. Each week, participants studied two hours from MOOCs and the other two hours in class, face to face. While students studied theoretical concepts from the MOOC, in face to face sessions they learned a general summary of the theory which was mainly applied to small example projects.

In this process, it was recommended that learners study some MOOC content every week during their own time and place. During this period some data collection techniques were used and in general students' satisfaction and learning improvement was observed.

In this hybrid MOOC application, some qualitative and quantitative data was collected with different methods. As a quantitative method, students completed some database diagram assignments during the term, and entered an exam. Moreover, they submitted small discussion post entries on homework each week about the given MOOC content. As qualitative data, very quick verbal surveys at the beginning of face to face meetings were done. The instructor asked some basic questions about the weeks' content like "how can you merge 2 tables with SQL's 'select' statement with alternative ways? Did you learn from the MOOC?" and he observed their level of understanding during the face to face learning. One advantage of hybrid MOOC application over a completely online MOOCs was to get weekly qualitative data from students. If they had some difficulty they could freely ask their instructor in meetings. So, it is a good idea to follow students' level of learning in the face to face meetings.

Moreover, in the hybrid MOOC application instructors can evaluate assignments and have a better understanding about what they learned. This is much more difficult in fully online MOOCs. In fully online MOOCs, the course is usually limited to content presentation and basic assignments. Some students may not learn the concepts although they completed a MOOC. It is because, usually in MOOCs they do not have to mirror their learning with projects, but in h-MOOC there is such an opportunity. In fully face to face courses, students need to take notes during class or instructor should share course content to allow them to

study on their own time. In this method even though students connect with the content, their effectiveness is limited compared to a recorded course. In MOOCs there is usually a recorded and carefully planned course content. Student could work by themselves with this and can repeat the content if they did not understand a part.

After three consecutive years of h-MOOC application, the course instructor -depending on his observations of the courses- perceived that students were learning subjects much more easily in the course hours as compared to fully face to face. For example, if the course instructor taught a new concept in a face to face section, because many students were learning the content for the first time, they had difficulties in understanding. But in h-MOOC sections, they performed clearly better because they already viewed new course content before the session that made them familiar with the concepts. Moreover, it is known that students learn at different speeds. Previously, although many students understood the course, some of them had problems and they asked questions during the classes. These questions were slowing down the learning process of the whole class in face to face sessions. But in h-MOOC, the instructor reminded them that they would repeat a part of the face to face course from MOOC if they did not understand. If they still had difficulties, there was an 'office hours' option as well.

According to instructor observations in the courses and small discussions with the students, the instructor perceived another good side of the h-MOOC application as having flexibility for both instructors and students. As students usually have complicated schedules, if they do not have to enter a course, it is observed that some prefer to do the work themselves rather than follow the course. This was especially valid if they could find some equivalent online courses like physics or computer science. In this manner, when they are given the opportunity to take some part of the courses on their free time, students were observed to be happy because of this flexibility. This was valid for the instructor. Because he assigned some online courses, he had additional time for managing the course quality and giving more feedback during office hours.

In fully online MOOCs, one problem seemed to be the lack of interaction (Daradoumis et al., 2013). It was also observed in our three years of study. But this was not a big problem in h-MOOCs because if students did not understand some part of the online form they could ask questions in the class. Actually, two types of students were observed: The first type of students prefer to search the content over the net and find answers to their questions on MOOCs. They seem to have high self-regulation skills. The second type of students prefer to ask the instructor questions instead of searching over the net. This type of students may not be happy

about using the MOOCs but only a few complaints were observed over the three years. We think that using h-MOOC can serve both types of students. In fully online MOOCs high dropout rates were observed (Liyanagunawardena et al., 2013). One reason would be that they were looking for extra interaction. On the other hand, although there is high interaction in fully face to face courses, some students hardly focus on the course, because they want to learn at their own pace. So, because h-MOOC type includes both methods together they would better help both groups of students.

6. Conclusion

In this chapter we mentioned that MOOC is a good type for a personal learning environment approach and also a good resource for data collection to improve learning outcomes. In addition, a PLE approach can be used in face to face learning with the help of MOOCs. So, a combination of MOOC and face to face sections that is called as hybrid MOOCs (PerezSanagustin et al., 2017) could be an effective way of learning.

In related literature it is understood that using a hybrid MOOC would result in mainly positive outcomes in terms of achievement and student feelings. Data collection is a critical part of the learning period because it will shape the overall process. Especially in the PLE approach, good data would be very helpful for students to allow them to shape their own learning process.

In h-MOOCs there is a need for special data collection approaches. It includes short interviews with students, observing their learning level, and evaluating homework. This data can help instructors to manage the course content. Actually, an instructor does not have to use only one MOOC resource (Bruff et al., 2013), instead he could take different chapters from different MOOCs and even alternative web resources. So, in time, the course content is getting better with the help of data. In this process, an instructor always needs to look at the effectiveness of the specific material that he used. In addition, an instructor should be encouraged to share this data in various platforms. In learning, there is a need for sharing learning experiences about such cases to make the learning process more effective.

Instructors can use LMS or similar software to manage their courses. It is good to assign the online courses, assignments, and discussion sections in a proper way. Many LMS, like Moodle or Edmodo, could be used for this aim. Instructors could share the required course content with URL to related sources. By using such LMS, students will have a live portfolio for their future life. Whenever they need, they could open required course from this portfolio.

Although it seems frustrating at the beginning, instructors could easily adopt this kind of application and over time it would become routine for them. Overall, because it gives some flexibility to both students and instructors, and it helps students to learn the content easier, this method makes learning more efficient. In addition, with the h-MOOC method, a course could be driven with more data and it has an advantage to better track the process. Of course, there could be problems as well. Some students, especially the ones that are more comfortable with the traditional education system and do not have high self-regulation skills may not like these methods. But, in the current generation, the number of this type of students is probably very limited. Finally, although the method works well in a Database Management Course, it may not be a proper way for learning in a different type of course. There is a need to research and evaluate the effectiveness of it in different courses like the social sciences, history, and the arts.

References

Banks, J. A., Au, K. H., Ball, A. F., Bell, P., Gordon, E. W., Gutiérrez, K., . . . Mahiri, J. (2007). Learning in and out of school in diverse environments: Life-long, life-wide, life-deep. Seattle: The LIFE Center and the Center for Multicultural Education, University of Washington.

Bruff, D. O., Fisher, D. H., McEwen, K. E., & Smith, B. E. (2013). Wrapping a MOOC: Student perceptions of an experiment in blended learning. Journal of Online Learning and Teaching, 9(2), 187.

Castano-Munoz, J., Kreijns, K., Kalz, M., & Punie, Y. (2017). Does digital competence and occupational setting influence MOOC participation? Evidence from a cross-course survey. Journal of Computing in Higher Education, 29(1), 28-46. doi:10.1007/s12528-016-9123-z

Daradoumis, T., Bassi, R., Xhafa, F., & Caballé, S. (2013). A review on massive e-learning (MOOC) design, delivery and assessment. Paper presented at the P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC), 2013 Eighth International Conference on.

Ebner, M., & Taraghi, B. (2010). Personal learning environment for higher education–a first prototype. Paper presented at the World conference on educational multimedia, hypermedia and telecommunications.

Hollands, F. M., & Tirthali, D. (2014). MOOCs: Expectations and Reality. Full Report. Online Submission.

Humanante-Ramos, P. R., García-Peñalvo, F. J., & Conde-González, M. Á. (2015). Personal learning environments and online classrooms: An experience with university students. IEEE Revista Iberoamericana De Tecnologias Del Aprendizaje, 10(1), 26-32.

Initiative, E. L. (2009). things you should know about… Personal Learning Environments. Educause..

Kloos, C. D., Muñoz-Merino, P. J., Alario-Hoyos, C., Ayres, I. E., & Fernández-Panadero, C. (2015). Mixing and blending MOOC Technologies with face-to-face pedagogies. Paper presented at the Global Engineering Education Conference (EDUCON), 2015 IEEE.

Konstan, J. A., Walker, J. D., Brooks, D. C., Brown, K., & Ekstrand, M. D. (2015). Teaching Recommender Systems at Large Scale: Evaluation and Lessons Learned from a Hybrid MOOC. Acm Transactions on Computer-Human Interaction, 22(2), 23. doi:10.1145/2728171

Liyanagunawardena, T. R., Adams, A. A., & Williams, S. A. (2013). MOOCs: A systematic study of the published literature 2008-2012. The International Review of Research in Open and Distributed Learning, 14(3), 202-227.

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2009). Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies. US Department of Education.

Moore, M. G. (1993). Theory of transactional distance. Theoretical principles of distance education, 1, 22-38.

Perez-Sanagustin, M., Hilliger, I., Alario-Hoyos, C., Kloos, C., & Rayyan, S. (2017). H-MOOC framework: reusing MOOCs for hybrid education. *Journal of Computing in Higher Education*, 29(1), 47-64. doi:10.1007/s12528-017-9133-5

Prensky, M. (2001). Digital natives, digital immigrants part 1. *On the horizon,* 9(5), 1-6.

Robinson, R. (2016). Delivering a medical school elective with massive open online course (MOOC) technology. *Peerj*, 4, 9. doi:10.7717/peerj.2343

Shah, D. (2015). MOOCs in 2015: Breaking Down the Numbers. Retrieved from https://www.edsurge.com/news/2015-12-28-moocs-in- 2015-breaking-down-the-numbers

Shearer, R., Gregg, A., Joo, K., & Graham, K. (2014). Transactional Distance in MOOCs: A critical analysis of dialogue, structure, and learner autonomy.

Siemens, G. (2013). Massive open online courses: Innovation in education. *Open educational resources: Innovation, research and practice*, 5, 5-15.

Swinnerton, B. J., Morris, N. P., Hotchkiss, S., & Pickering, J. D. (2017). The Integration of an Anatomy Massive Open Online Course (MOOC) into a Medical Anatomy Curriculum. *Anatomical Sciences Education*, 10(1), 53-67. doi:10.1002/ase.1625

Tsui, E., & Sabetzadeh, F. (2014). Lessons Learnt from and Sustainability of Adopting a Personal Learning Environment & Network (Ple&N). *International Association for Development of the Information Society*.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective.