

University POLITEHNICA of Bucharest

Faculty of Automatic Control and Computers, Department of Computers

## Habilitation Thesis

---

in Computers, Information Technology and Systems Engineering

## Automated Assessment of Comprehension and Collaboration in Individual and Social Learning Contexts

---

presented by Mihai DASCĂLU

May 2017, Bucharest, Romania



# Abstract

---

In conjunction with the increasing popularity of e-learning technologies, the need to automatically evaluate the two closely related activities of individual learning (consisting both of reading and of writing texts) and collaborative learning has increased significantly. A deeper understanding of discourse is needed at the level of these assessments, even if a superficial or surface analysis is easy to achieve. In this context, we introduced Cohesion Network Analysis and refined the polyphonic model of the discourse, originally described by Ștefan Trăușan-Matu and inspired from Bahtin's studies on dialogism. These models enable us to process both general texts and online conversations, as well as to integrate specific comprehension and collaboration activities into a single platform – *ReaderBench*.

This habilitation thesis presents my scientific, professional and academic achievements in our expanded research group, all of which are based on interdisciplinary experiments, covering the domains of Computer Science, Educational Psychology, Computational Linguistics, and Philosophy. In short, I was the head of my class in 2009 with a GPA of 10/10 at Politehnica University of Bucharest (UPB), and I have a double PhD in double coordination with the highest distinctions in Computers and Information Technology ("excellent" mention, UPB), and Educational Sciences (mention "Très honorable avec Félicitations", Grenoble-Alpes University, France); my PhD thesis was afterwards published as a book in Springer, *Studies in Computational Intelligence*. So far I have accumulated extensive experience in national and international research projects (PC D HUB-TECH, POC G NETIO, H2020 RAGE, ERASMUS + ENea-SEA, FP7 LTfLL, FP7 ERRIC, and CNCSIS K-teams), with over 130 published papers, including top conferences (AAAI, CogSci, AIED, STIs, CSCL), renowned international conferences (ICALT, EC-TEL, ICWL, ISPDC, AIMS), and prestigious journals (Elsevier *Computers in Human Behavior*), which brought me over 700 citations.

Currently I am an associate professor at UPB, responsible for the courses regarding Object-Oriented Programming, Semantic Web Applications, and Data Mining and Data Warehousing, as well as an associate member of the Romanian Academy of Sciences. Complementary to my skills in Natural Language Processing, discourse analysis and Computer Supported Collaborative Learning, I hold multitude professional certifications (e.g. PMP, PMI-RMP, PMI-ACP, CBAP, CISA, C|EH, CISSP, and OCMJD), as well as a wide experience in strategic projects relying on non-refundable funds from the European Union, World Bank, or USTDA. In 2013, I received the distinction "IN TEMPORE OPPORTUNO"

for the most promising young researcher at UPB, and in 2015 I obtained a Senior Fulbright scholarship that opened long-term research collaborations.

As a specificity of the performed analyses addressing the individual learning process, we focused on creating a multidimensional model of textual complexity that integrates surface, word, morphological, syntactic, semantics and discourse structure indices. From a complementary view, collaborative learning, i.e. the social dimension, focuses on assessing the involvement of participants and of collaboration based on two computational models: a) a model centered on social knowledge building and derived from Cohesion Network Analysis, as well as b) a polyphonic model derived from the inter-animation of voices.

Our approach integrates advanced Natural Language Processing techniques and focuses on providing both qualitative and quantitative estimates of the learning process. Various use cases are presented throughout the thesis, arguing that our system has undergone intense validations, and indicating its ability to simulate human expert ratings. Starting from the undergone experiments, we succeeded in extending the perspective of collaboration with regards to achieving a coherent representation of discourse through the inter-animating of voices pertaining to different participants, as well as through text cohesion. Thus, one of the most important objectives of our models is to enhance comprehension as a "mediator of learning", by providing automated feedback to both learners and tutors. The main advantages achieved through the implementation of the *ReaderBench* framework are flexibility, extensibility, as well as specificity to cover multiple processing stages and educational experiments.

I strongly believe that automated models of comprehension will continue to be used more and more intensively in educational processes. My priorities consist of maintaining and expanding our research network, while further enhancing our *ReaderBench* framework. Moreover, I hope and wish that my current career development, as well as current research prospects, will create the prerequisites for new projects, collaborations and publications.