2016

A Formalism For PLAN – A Big Data Personal Learning Assistant For University Students

Timothy Arndt
Cleveland State University, t.arndt@csuohio.edu

Angela Guercio
Kent State University - Stark, aguercio@kent.edu

Publisher's Statement
This article was published in the Journal of e-Learning and Knowledge Society

Original Published Citation
Arndt, T. & Guercio, A. (2016). A formalism for PLAN - a big data personal learning assistant for university students. Journal of e-Learning and Knowledge Society, 12(2), retrieved from http://www.je-lks.org/.

Je-LKS
Journal of e-Learning and Knowledge Society
The Italian e-Learning Association Journal
INVITED PAPERS
BIG DATA, COGNITIVE COMPUTING AND INNOVATIVE TEACHING MODELS
Vol. 12, n.2, 2016
ISSN: 1826-6223 | eISSN: 1971-8829

A FORMALISM FOR PLAN A BIG DATA PERSONAL LEARNING ASSISTANT FOR UNIVERSITY STUDENTS

Timothy Arndt[1]
Angela Guercio[2]

[1]Department of Information Systems - Cleveland State University, Cleveland, OH, USA - t.arndt@csuohio.edu
[2]Department of Computer Science - Kent State University at Stark, Canton, OH, USA - aguercio@kent.edu

Keywords: personal learning system, big data, learning recommendation systems, learning models, e-learning.

Big Data-based methods of learning analytics are increasingly relied on by institutions of higher learning in order to increase student retention by identifying at risk students who are in need of an intervention to allow them to continue on in their educational endeavors. It is well known that e-Learning students are even more at risk of failing out of university than are traditional students, so Big Data learning analytics are even more appropriate in this context. In this paper, we present our approach to this problem. We wish to place control of a student's learning process in his own hands, rather than that of the learning institution in order to decouple the student from the institution since the goals and motivations of these two may not be completely aligned. In this way, we empower the student by giving him control of the personal learning system which employs Big Data techniques to generate recommendations on how to reach a set of learner-specific learning goals. We present the formalism which underlies our system, the

for citations:
Arndt T., Guercio A. (2016), *A formalism for PLAN – a Big Data Personal Learning Assistant for university students,* Journal of e-Learning and Knowledge Society, v.12, n.2, 13-25. ISSN: 1826-6223, e-ISSN:1971-8829

architecture which implements the system, scenarios for system use, a survey of related works and thoughts on how the system will be implemented in a prototype in the future.

1 Introduction

The rise of e-Learning as a viable option for post-secondary education promises to open up university education to a wider range of students than ever before. However, just as in traditional, face-to-face learning guidance is required for students to be able to achieve to the best of their abilities. Traditional universities are increasingly paying attention to student retention rates and using Big Data techniques in learning analytics to identify those students who are most at risk so that an intervention can be scheduled to best allow the student to succeed. It is well known that the retention problem is multiplied in the e-Learning environment, so learning analytics should prove to be even more useful in this context.

In this paper, we describe our approach to the use of Big Data to the specific problems of e-Learning students. Our approach is unique in that is personalized for individual students, and even more importantly the approach is under the control of the student – it is not controlled by a learning institution whose goals and motivations may not align completely with those of a student. In this way, the student can choose the institution which best aligns with his learning goals, select the programs and instructors within those institutions which best help reach those goals, and can take this system with him as he (possibly) moves from one institution to another.

We present a formalism for our personal learning assistant. This formalism integrates Big Data (learning and social network) analytics, a finite state transducer which acts as a recommender system for the learner, the learner's calendars, location, a learner profile, etc. into a complete system for assisting the student to his personalized goals. The aggressiveness of the recommendations can be controlled by the user. We next lay out several scenarios of the system in action which help to explain how it will be used in practice. The cloud-based architecture of the system with a mobile app user interface is then described. Finally, we survey related research and then give conclusions pointing out how we will implement the system and other future research goals.

2 State of the Art

The design a Personal Assistant that helps people to handle daily tasks is something that continues to interest modern research. The term personal assistant has been used in research to indicate a software entity that assists either a human or another software entity in some type of activity. As Huhns and

Singh state in (Huhns, 1998) "*a personal assistant differs from a personalized search program or e-mail filter in that it is inherently network-based, interactive, and adaptive. Personal assistant applications are also general-purpose and long-running.*"

The studies that propose the design and development of a personal assistant can be classified in two categories: those which study personal assistants that provide software support to a human daily work; and the other which study personal assistants that sustain and improve cooperation between software agents. Examples of the second category can be found in (Chen, 2007) where the authors focus their research to a memory mechanism to enhance the intelligence of a personal assistant agent while interacting with other agents. Fan *et al.* (2008) instead focus their work on the personal assistant for reducing the cost of communication in CPN, a coordination mechanism in multi-agent systems.

Since in this paper we explore a formalism of a personal assistant whose service beneficiary is a human being (i.e. a student), we overview primarily the research for the design and development of a software personal assistant that supports human activities.

The applications areas of a personal assistant also vary greatly: from transportation, to medical, to education, etc.. According to the application area of the personal assistant and the problem at hand some formalism has been suggested. However, there is no existing formalism proposed in the area of education for a learning assistant that uses big data methods, that is personalized for individual students, and is under the control of the student, which is the goal of our research. For example, in the area of transportation, a scientific study that uses stochastic methods to model the personal preferences of a PA in the area of urban transportation is described in (Bell, 1997; Sheffi, 1985). Two alternative approaches, in the same application area, use rational decision algorithms (Brooks, 2002; Robbins, 2007) defining the planning process as a multistep process which covers the stages from problem identification to the solution. None of those formalisms takes into account the problem of big data.

Pozna *et al.* (2013) provide formalism for a Personal Assistant who is in charge of planning daily activities of a human operator. The model is based on the rational decision algorithm and uses a mathematical model to generate solutions to the problem, solutions that are refined with the user's feedback. Contrary to our approach this model bases its solution on a list of a known prior list of actions which has an initial time and an ending time. In our case the list of actions is not known prior, it evolves when the learning analytics discover new data and the user decides or not to act upon.

In the area of education, all of the personal assistants we are aware of, are designed and customized for an entity or an organization to provide better support for the student and replace and/or to assist the work of an advisor or of an instructor. In (Lonn, 2014) learning analytics are used to support advisors in identifying student's academic performance; in (Nam, 2014; Siegfried, 2003) the personal assistant supports the student in selecting the courses in which to enroll, and in what combinations, to minimize the impact each student's bad choice has for academic success; in (Arnold, 2012) the software system helps instructors to employ the power of learner analytics to provide real-time feedback to a student to increase student success. Similarly, Vailardi et al in (Vailardi, 2009) present an advising system that uses data mining to support a student in its chosen academic itinerary at the Universidad de Lima. Zhang et al (Zhang *et al.*, 2014) design a personal assistant robot platform for students and employee to help to manage their learning, life and work better, in the fields of curriculum management, diary management, financing management and chatting system. The system uses a rule library and artificial intelligent methods for rule learning to derive assistance messages passed to the user via an instant messaging system. In (Sheehanm, 2012) and (d'Aquin, 2013) the authors mine student scholastic data at their institution to improve the academic experience for their students.

In all these cases, the personal assistant is designed by following primarily the requirements of the institution to which the student belongs or of an institution in general. The personal assistant we present is designed by the student for the student, created and customized by the student according to his or her own learning goals.

In some cases, personal assistants are used in connection with prediction systems. In those cases big data and analytics are used to foresee possible students critical situations, such as predicting grades that may hurt the student GPA (Willis III *et al.*, 2013), or predicting student performance in a classical (Thai-Nghe, 2011) or in an online course environment (Tanner, 2010). In those personal assistants, prediction is used to guide recommendation.

Recommendation systems (Jannach, 2010) have become an integral part of the Web experience today in connection with big data and learning analytics. A personal assistant for the student success incorporates a recommendation system that uses a diverse set of data sources that go from the social network, to the user profile, to the network in general to provide suggestions, advices, and alerts. The system presented in (Farzan, 2006) describes a community-based hypermedia recommendation system that employs social navigation and community of learning to tackle the problem of information overload created by the large amount of data and courses available on the net, and extracts community wisdom derived from explicit and implicit user feedback (from the user interactions with the system) to identify proper recommendations. Another recommendation system, called Sherpa, is described in (Bramucci, 2012). Sherpa is a software advisor that offers personalized recommendations for courses, information and services and utilizes both human and machine intelligence. However, as other systems mentioned above, Sherpa is designed for an institution.

3 A formalism for PLAN

In this section we will describe the formalism which underlies our personal learning assistant PLAN (Personal Learning AssistaNt). The formalism serves as documentation for the structure of the system, will guide the development of future prototype systems, and can be used to prove properties of the system as well.

A personal learning assistant supports a user in achieving desired learning goals. The formal definition of a personal learning assistant is described in this section.

A personal learning assistant uses multiple source of information to provide appropriate recommendations and alerts when active. Some data are collected through the direct interaction with the user, while other data are mined by searching across available data sources. The first source of information used by the personal learning assistant comes from the user's learning profile.

Definition 1 A user's *learner profile* P is an n-tuple $P = (a_1, a_2, \ldots, a_n)$ where each a_i is an attribute and $a_i \in D_i, 1 \leq i \leq n$, the domain of the i-th element of P.

An additional source of information is provided by the user's social network from which desired data are extracted.

Definition 2 A user's *social networks* SN is a set $SN = \{G_1, G_2, \ldots, G_n\}$ where each $G_i = (N_i, L_i)$ is a graph, $1 \leq i \leq n$, where N_i is a set of nodes and $L_i \subseteq \{\{N_{ij}, N_{ik}\} | N_{ij}, N_{ik} \in N_i\}$ is a set of links. Nodes and links are further described by a set of attributes, where the number and type of attributes is dependent on the particular social network.

Finally, additional data are mined from the multiple data sources available across the networks.

Definition 3 A user's *data sources* DS is a set $DS = \{DB_1, DB_2, \ldots, DB_n\}$ where each DB_i, $1 \leq i \leq n$, is a set is a set of object $DB_{ij} \in DB_i$ of the type (Key, Value). Standard operations GET(Key, DB_i), SET(Key, Value, DB_i), and DELETE(Key, DB_i) are defined on the data sources which are key-value databases to be mined by analytics methods to help reach the desired goals.

Mining significant data from data sources means to be able to identify data that can be of interest to the user. So while an event can be considered interesting for achieving a learning goal, that event could be irrelevant if the user is unable to participate in it. This means that there are constraints that should be taken into account in order to select significant data and provide useful recommendations. The two primary constraints come from the spatial and temporal information associated with the user. The temporal information is derived from a calendar. The spatial information is derived from the user's geolocation which, from the implementation point of view, is associated with the GPS coordinates extracted from the sensor of the device where the personal assistant has been launched from.

Definition 4 A user's *calendars* CAL is a set CAL = $\{CAL_1, CAL_2, \ldots, CAL_n\}$ where each CAL_i, $1 \leq i \leq n$, is defined as a mapping c: Year × Month × Day × Hour → Event; i.e. a mapping from a date and time to events.

Definition 5 A user's current *location information* LOC is a pair (latitude, longitude) where the values have their common geographic interpretation.

The amount of information produced by the personal learning assistant could be very large and overwhelm the user. A level of aggressiveness must be used to customize the personal learning assistant to a user-desired level.

Definition 6 A user's chosen *aggressiveness* of recommendations AGG is a real-valued number $0.0 \leq AGG \leq 1.0$.

An AGG of 0.0 will result in no recommendations being generated by the personal learning assistant, while 1.0 will result in a maximum number of recommendations being generated.

The activity of the personal learning assistant PLAN can be formally described by a learning finite state transducer as stated in the following definition.

Definition 7 A user's *learning finite state transducer* LFST is a finite state machine represented as the six-tuple LFST = $(\Sigma, \Gamma, S, s_0, \delta, \omega)$, where Σ is a finite input set of derived knowledge items, Γ is a finite set of output knowledge items, S is a finite non-empty set of learning states, δ is a state transition partial function defined as a mapping $\delta: (\Sigma \cup \{\varepsilon\}) \times S \rightarrow S$ where ε is the empty knowledge item, ω is an output function defined as a mapping $\omega: (\Gamma \cup \{\varepsilon\}) \times S \rightarrow S$ where ε is the empty recommendation, and s_0 is the initial learning state ($s_0 \in S$).

The LFST moves from state to state in order to reach the desired learning goals and in each transition produces 0 or more outputs.

Definition 8 A user's learning goals LG is a set LG $\subseteq$ S.

Definition 9 (PLAN - Personalized Learning AssistaNt) A *personalized learning assistant* is defined as an eight-tuple PLAN = (P, SN, DS, CAL, LOC, AGG, LG, LFST).

When the system is initialized for a user, the PLAN generates the user's learner profile P based on an interactive process with the user, as well as on the basis of a default profile. The learner profile may be refined as the learning process advances. The system also interrogates the user to determine the user's set of learning goals LG. The learner profile and learning goals are then used by the system to generate the LFST. The learner identifies his calendars, relevant data sources, and social networks which are then imported into the system and become part of his PLAN. Default data stores and calendars may also be imported into the system (we don't assume that the user is an expert on available DS).

Data is mined from the DS, which are key-value data stores, as well as from the user's social networks SN, using traditional data mining processes. This results in knowledge items being discovered as data mining proceeds and as the DS and SN are continuously updated. Each time a knowledge item is discovered, the state transition function δ may result in a transition to a new state in the LFST (but not every knowledge item results in a transition since δ is a partial function). Calendar events from the user's calendar set CAL may also be generated as time passes. These calendar events are treated as knowledge items by the LFST.

A state transition generally results in zero or more recommendations (i.e. the output function of the LFST) being made to the learner (e.g. to take a section of a particular source). On the other hand, some action of the learner may result in a state transition (e.g. the user successfully completing a course). In this case, the completion of the course is treated as a new knowledge item, which results in a state transition, but no recommendation to the learner is required. The state transition function depends on both AGG and LOC. The user's location may affect the state transition, thus resulting in a location-specific recommendation to the learner (e.g. take a lower-level course at the local community college). The AGG setting also affects the state transition function in the following way. Each learning state in the LFST has an attribute, which is an estimate of how far away the state is from a learning goal state.

In general, a transition will only take place from one learning state to another, if the transition takes the learner to a state which is closer to one of his learning goals. The AGG setting determines *how much closer* the new learning state has to be to a goal in order for the transition to take place and the recommendation to the user generated by the output function. If AGG is set to 0, *only* transitions which take the user to a learning goal state are taken. If AGG is set to 1, all transitions which take the user closer to a learning goal are taken. The AGG setting is a personalization feature, which allows the learner to control how aggressive the PLAN is in making recommendations – it permits the user to avoid getting overwhelmed by recommendations.

4 Architecture

In this section we will describe the architecture of the personal learning assistant system. The architecture is shown graphically in figure 1.

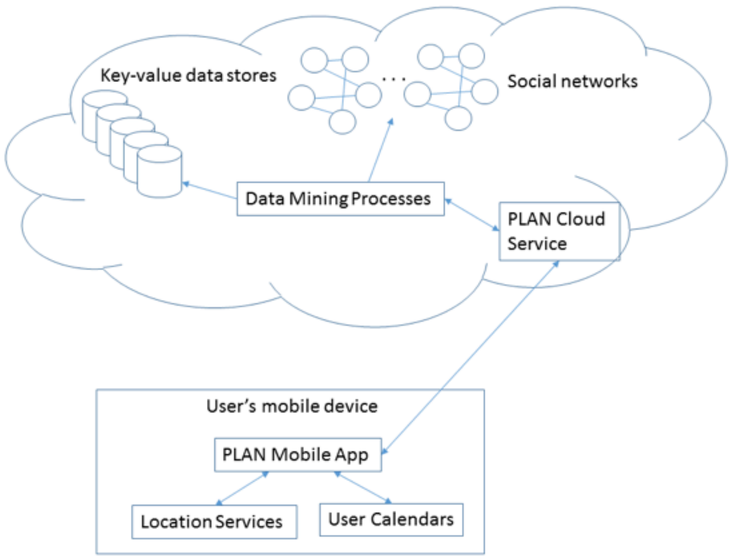

Fig. 1 - The high-level architecture of the personal learning assistant system.

The high-level description of the architecture shown in figure 1 indicates that the system is structured as a mobile app which communicates with a cloud service which performs the main share of the computation. The PLAN Cloud Service interacts with standard data mining processes which run in the cloud and which work on two categories of data: the key-value data stores and the user's social networks. The key-value data stores represent the raw matter used for learning analytics. Sources of the data which can be analyzed include institutional data about students, courses, applicants, as well as a particularly rich field to mine for data - that associated with online courses and Course Management Systems (CMS). Learning analytics uses a combination of institutional data, statistical analysis, and predictive modeling to create insight which can be used to develop a strategic plan for enhancing academic outcomes. We will incorporate standard learning analytics techniques, as well as more specialized ones which are enabled by our approach. The particular form of the specialized techniques is part of our future research. To start with, our prototype system will likely use standard approaches and modules. One issue that arises immediately is – where will the information that will be the input to the analytics process come from? In institutions of higher educations, this is not a problem, since the institutions own the data that they generate. The situation is different in this case however, since the students do not generate or own the data that is needed for the personal learning assistant. The institutions of higher education could make this information publicly available so that it could be used by students (after it has been suitably scrubbed to make sure that privacy concerns are met). If the institutions are unwilling to make the information available, they may need to be encouraged to do so (by government agencies in the case of publicly-funded universities, by donors or accreditation boards in the case of private institutions). Some of this information is already publicly available (sometimes due to government regulations) either individually at the institutions, or collected by agencies or commercial entities (Arndt, 2016). This learning analytics data is stored in key-value database stores after a minimal amount of processing.

In addition to this data, the data mining processes interact with the student's social networks – both online social networks such as Facebook and Twitter, as well as more informal social networks such as those identified by email and text message communication as well as those which are deduced by examining the course rosters of the courses the students are enrolled in. These social networks form another important asset in the student's learning process, being sources of expertise in course topics, various university processes, job markets and so on. The data mining processes can interact with online social networks through the APIs that they provide and through the more informal social networks through bespoke processes which may be developed.

The data mining processes generate the knowledge items which the PLAN Cloud Service uses to navigate the user's personal LFST. The Cloud Service generates recommendations as it moves from state to state in the personal LFST, and these recommendations are pushed to the PLAN Mobile App to be presented to the learner. The PLAN Mobile App constantly informs the PLAN Cloud Service of changes to the student's location and passes calendar events from the student's calendars, which are generally held on the mobile device to the PLAN Cloud Service. User interactions and interrogations also take place via the PLAN Mobile App.

5 Scenarios

In this section we give two scenarios of how we expect the personal learning assistant to function in practice. These scenarios are not meant to be an exhaustive description of the capabilities we foresee for PLAN, but merely to give an idea of how it might be used by university students to help them pursue their learning goals.

Scenario 1 (initialization process)

Sue is interested in pursuing a career in Information Technology, so in initializing her personal learning assistant she includes earning a degree in a subject which will help her enter an IT field as one of her career goals. She indicates as part of the interrogation process which accompanies initialization that she prefers to attend college near home, full-time and that she needs to be cost-conscious in choosing an institution. This information forms part of her learner profile. She uploads her electronic learning record (high school and test records) which is added to the standard key-value data stores after some processing and imports her social networks as well. On the basis of Sue's preferences, the system performs data mining on the analytics data and suggests several local institutions that would be suitable for Sue's education, and several majors at the institutions (found on the basis of hiring patterns, Sue's academic strengths and interests, etc.). After Sue makes her initial choice, the LFST is generated and Sue is ready to start her academic career.

Scenario 2 (guidance through the course scheduling process)

Sue has completed two years of courses in her chosen field of technical writing at the local state university, and is deciding which courses to register for in the coming year. Because she feels the need for additional help in this process, she increases the AGG setting of her PLAN so that it will generate recommendations more aggressively. The state that the LFST is in reflects the fact that Sue has completed the first two year of courses and needs to register for the upcoming year (among many other things). Over the past two years, Sue's learner profile has been refined, reflecting the experiences of those years. Based on this profile (among other things), when Sue increases the AGG setting the LFST generates a recommendation that Sue take a set of general education courses at the local community college due to their lower cost and the fact that the data mining process has determined that students taking courses at this community college are not adversely affected in their subsequent academic work at Sue's state university. This recommendation to take courses away from her current institution even though equivalent ones are available locally shows that the interests of the institution and the learner are not always aligned – thus validating our student-controlled (rather than institution-controlled) approach. The PLAN recommends a schedule of courses with those instructors which analytics show are most effective in terms of student outcomes. The PLAN also recommends traditional face-to-face courses rather than online ones since Sue's learner profile (based on her results in the previous two years) reflects the fact that Sue does better in those courses. Once the recommendation is made (output by the LFST), the LFST moves to a new state. When Sue accepts the recommendation and registers for the courses at the community college, this results in a knowledge item being input to the system and the system transitions to a new state (a "transient student at local community" state). No recommendation is made (output generated) by this transition. Once in this state, the system recognizes that Sue has strong ties in her social networks with a current student at the community college and generates a recommendation that Sue consult her friend about logistical issues at the community college (transportation, parking, dining, etc.).

Conclusions and Future Research

In this paper we have presented a formalism for PLAN a Personal Learning Assistant for university students which empowers students by allowing them to control their university education. Scenarios for the use of PLAN have also been presented. This research is distinguished from other efforts in this area in that Big Data learning analytics are employed by the learner, rather than the institution, allowing the learner to pursue his goals wherever they may lead him – even to a different institution of higher learning if that would be more efficient in meeting his goals. The approach is also distinguished by having an underlying formalism. The advantage of this is that it serves as a precise specification of the system and may allow us to formally prove properties of PLAN.

We are currently planning a prototype implementation of PLAN. We are identifying default learning analytic data stores, data mining engines, as well as templates for the LFST and learner profile which will be the starting point for the interrogation process described in scenario 1 above. We will be conducting usability studies with university students at Cleveland State University and Kent State University and determining how the students' social networks can be integrated into the system. We foresee that the experiences gained through the prototype development and usability study will feed back into the design of PLAN and lead to modifications to improve usability in practice.

We are also exploring the possibility of proving properties of PLAN, for example reachability of learning goals, in order to exploit the formalism presented in this paper. All of these developments will be documented in future papers.

REFERENCES

Arndt T., Guercio A. (2016), *Student-centered analytics for postsecondary education*, in: Competencies, Challenges and Changes in Teaching, Learning and Educational Leadership in the Digital Age, J. M. Spector, D. Ifenthaler, D. Sampson and P. Isaias (eds.), Springer.

Arnold K. E., Pistilli M. D. (2012), *Course signals at Purdue: using learning analytics to increase student success*, in: Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK 2012), Vancouver, BC, Canada, 267-270.

Bell M. (1997), *Transportation network analysis*, John Wiley& Sons, Chichester.

Bramucci R., Gaston J. (2012), *Sherpa: increasing student success with a recommendation engine*, in: Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK ‘12), Simon Buckingham Shum, Dragan Gasevic, and Rebecca Ferguson (eds.), ACM, New York, NY, USA, 82-83, doi:10.1145/2330601.2330625 http://doi.acm.org/10.1145/2330601.2330625.

Brooks M.P. (2002), *Planning theory for practitioners*, Chicago: American Planning Association, 175-6.

Chen K. J., Barthes J. P. (2007), *Enhancing intelligence of personal assistant agent using memory mechanism*, in: Proceedings of the 11th International Conference on Computer Supported Cooperative Work in Design (CSCWD 2007), Melbourne, Vic., 361-365, doi: 10.1109/CSCWD.2007.4281462.

d’Aquin M., Jay, N. (2013), *Interpreting data mining results with linked data for learning analytics: Motivation, case study and directions*, in: Proceedings of the 3rd International Conference on Learning Analytics and Knowledge (LAK 2013), Leuven, Belgium, 155-164.

Fan G., Huang H. & Jin S. (2008), *An extended the contract net protocol based on the personal assistant*, in: Proceedings of ISECS International Colloquium on Computing, Communication, Control, and Management (CCCM ‘08), Guangzhou, 603-607, doi: 10.1109/CCCM.2008.199.

Farzan R., Brusilovsky P. (2006), *Social navigation support in a course recommendation system*, in: Adaptive Hypermedia and Adaptive Web-Based Systems, 91-100, Springer Berlin Heidelberg.

Huhns M. N., Singh M. P. (1998), *Personal assistants*, IEEE Internet Computing, 2(5), 90-92, Sep/Oct 1998, doi: 10.1109/4236.722236.

Jannach D., Zanker M., Felfernig A. & Friedrich G. (2010), *Recommender systems: an introduction*, Cambridge University Press.

Lonn S., Teasley S. D. (2014), *Student explorer: a tool for supporting academic advising at scale*, in: Proceedings of the 1st ACM conference on Learning @ scale conference (L@S ‘14), ACM, New York, NY, USA, 175-176, doi:10.1145/2556325.2567867 http://doi.acm.org/10.1145/2556325.2567867.

Nam S., Lonn S., Brown T., Davis C.-S. & Koch D. (2014), *Customized course advising: investigating engineering student success with incoming profiles and patterns of concurrent course enrollment*, in: Proceedings of the 4th International Conference on Learning Analytics And Knowledge (LAK '14), ACM, New York, NY, USA, 16-25, doi:http://dx.doi.org/10.1145/2567574.2567589.

Pozna C., Foldesi P. & Kovacs J. (2013), *The personal assistant application, problem definition*, in: Proceedings of IEEE 4th International Conference on Cognitive Infocommunications (CogInfoCom 2013), Budapest, 851-856, doi: 10.1109/CogInfoCom.2013.6719215.

Robbins S.P., Judge T.A. (2007), *Organizational behavior* (12th ed.), Upper Saddle River, New Jersey, Pearson Prentice Hall, 156–8, ISBN 978-0132431569.

Sheffi Y. (1985), Urban transportation networks, Prentice Hall Englewood, New Jersey.

Sheehan M., Park Y. (2012), *pGPA: a personalized grade prediction tool to aid student success*, in: Proceedings of the sixth ACM conference on Recommender systems (RecSys '12) ACM, New York, NY, USA, 309-310, doi:http://dx.doi.org/10.1145/2365952.2366027.

Siegfried R. M., Wittenstein A. M. & Sharma T, (2003), *An automated advising system for course selection and scheduling*, J. Comput. Sci. Coll., 18(3), (February 2003), 17-25.

Tanner T., Toivonen H. (2010), *Predicting and preventing student failure using the k-nearest neighbour method to predict student performance in an online course environment*, IJLT, 5(4), 356 – 377.

Thai-Nghe N., Horvath T. & Schmidt-Thieme L. (2011), *Factorization models for forecasting student performance*, Educational Data Mining (Eindhoven, Netherlands, 2011), 11-20.

Vialardi C., Bravo J., Shafti L. & Ortigosa A. (2009), *Recommendation in higher education using data mining techniques*, Educational Data Mining, Cordoba, Spain, 190 – 199.

Willis III J. E., Campbell J. P. & Pistilli M. D. (2013), *Ethics, big data, and analytics: A model for application*, EDUCAUSE Review, available: http://www.educause.edu/ero/article/ethics-big-data-and-analytics-model-application

Zhang Y., Hu Y., Zhang P. & Zhang W. (2014), *Design of personal assistant robot with interactive interface on learning and working*, in: Proceedings of the 5th International Conference on Intelligent Systems, Modelling and Simulation (ISMS), Langkawi, 502-506, doi: 10.1109/ISMS.2014.92.