Doctor of Professional Studies in Computing: A Categorization of Applied Industry Research

Lisa R. Ellrodt, Ion C. Freeman, Ashley J. Haigler, and Suzanna E. Schmeelk

Seidenberg School of Computer Science and Information Systems

Pace University

New York, New York

***Abstract*—This is a full paper in the Innovate Practice category examining a Doctor of Professional Studies (D.P.S.) in Computing program at Pace University which is a specialized degree program designed for active Information Technology (IT) professionals with at least five years of full time experience in the computing field. The first question that guided our research is how is the Pace University D.P.S. program related to other doctoral computing programs in the United States and globally. The second question that guided our inquiry is dissertation topics pursued by graduates in dissertation research, specifically by IT professionals in the Pace University program. To answer this second question, we analyzed the first 114 dissertation abstracts that have been defended at Pace University in the Seidenberg School of Computer Science and Information Systems. We used machine learning and natural language processing to determine commonalities among research topics in order to gain an understanding of the topic categories and topic spread produced by the program.**

Keywords— Doctoral Programs, Full time working IT Professionals, Computational Thinking, Computing, Education, Natural Language Processing, Machine Learning

I. Introduction

The Doctor of Professional Studies (D.P.S.) in Computing Studies is a specialized degree program designed for active IT professionals with at least five years of full time experience in the computing field. This allows for a student to practice computing while earning their degree. The part-time program allows working IT professionals, managers, and executives to continue their computing education within a three year weekend-oriented program. The classes follow a hybrid design: monthly meetings occur in person in Pleasantville, New York; and, weekly meetings occur at least two times a week online within smaller team groups comprised of both faculty and students. This design creates a collaborative learning experience due to to the diverse backgrounds of the team members. The nature of the team cohort is strengthened by the bonds formed between the team members which lends support to each student as they complete this rigorous research based program as well as focus on the daily stresses of work life[11].

Pace faculty—who come to Pace from Cornell, IBM, Oracle, West Point, and other well known full time IT careers—work with students during evening and weekend hours enabling working professionals to continue to succeed in their professions (e.g., meetings, travel, promotions, etc.) while enhancing and expanding their chosen domains of IT expertise. Students are enabled to work on their dissertations during their off work hours and on average defend their dissertation after their third year of studies. Many of the IT professionals in the program write a dissertation that is correlate to their full time career.

Our conceptual framework builds on research into adult learning. Research has shown that most adult Americans are lifelong learners. A Pew Research Center report [23] on adult learning found that over half of all adult Americans have invested resources (e.g., time, money, etc.) into "additional training in the past 12 months to improve their job skills or expertise connected to career advancement."

Our research is twofold. First, our research explores how the Pace University D.P.S. program relates to other doctoral computing programs in the United States and globally. Merritt et al. [10] assessed the D.P.S. program after five years. Along with a five thrust assessment analysis reported by Merritt et al., their paper introduced dissertation topics that were being defended within the program within the first five years of the program. This research extends the work of Merritt et al. [10] within a second research question where we analyze all defended topics pursued by graduates in dissertation research. After working to categorize the dissertations by hand, we found that our opinions on categories were individual to the reviewer; standardization was difficult. As such, we employed machine learning and natural language processing to categorize the first 114 dissertation abstracts that have been defended at Pace University in the Seidenberg School of Computer Science and Information Systems. Our paper reports on topics defended by doctoral students to gain an understanding of the research needs facing IT working professionals in the program.

II. Doctoral program research

There is surprisingly only a few scholarly research studies relating to doctoral dissertations or doctoral degree programs in Computing, Computer Science, and Information Systems. What research that exists on doctoral degree programs, and is

978-1-5386-1174-6/18/$31.00 ©2018 IEEE

present on either the ACM Digital Library or the IEEE Xplore web portals, tends to be polarized as either related to program curriculums or student demographics. We summarize the research, as given by searches on these web services, in the following two subsections.

A. Research on Doctoral Program Curriculums

There are a few studies on doctoral program curriculums across the globe. The research is claiming to find insights on program curriculums as informed by interdisciplinary, industrial, and student needs.

Moreno [9] reports on a work in progress study where they believe that "the circumstances in which postgraduate programs are offered have a major impact on their quality." Their one page research progress reports that their research plans to provide a better understanding of the process, content and outcome of different programs' models of doctoral education in Computer Science. The study will examine the differences between Anglo-American doctoral degrees and traditional European Master-apprentice model degrees.

Bernat and Grimson [1] proposed improvements to doctoral program ranking strategies. Their findings included two different ranking criteria, one based on surveys and one based on regression.. The two ranking approaches yield different results. Part of their improvement plan includes using other metrics of importance, besides the traditional ranking system metrics, in a new ranking system.

Research by Hoganson and Amoroso [6] reported on the design of a an interdisciplinary doctoral program between Information Systems and Computer Science to combine the strengths of the two related disciplines. According to the paper, the new program design attempts to capitalize on areas of overlapping expertise to meet the needs of high-tech employers and innovators.

Joyce and Young [7] describes the processes involved in developing and implementing the Doctor of Computing (DComp) programme at UNITEC Institute of Technology in New Zealand. The DComp program aimed to provide extended education to working IT professionals of over 10 years. Their research found that the greatest DComp challenge was in meeting the students' wide range of interests and facilitating the integration of the many and varied inputs.

Research by García-Peñalvo [3] showed that creating a new PhD computing interdisciplinary program in Spain met growing needs across different industry sectors.

Research from Heldal et al. [5], in Sweden, developed an instrument to support systematic communication and collaboration between the different stakeholders involved in industrial based doctoral projects.

B. Research on Doctoral Degree Student Demographics

The ACM and IEEE only list a few studies related to doctoral degree student demographics. These studies tend to analyze the following: age, enrollment status, or gender.

In a Turkish Engineering student body case study analysis published by Gulluoglu [4], the research showed student statistics on age, gender and time enrollments within the Engineering school advanced degree programs, which included doctoral programs. The quantitative study showed that 45% of the students were under 30 years old, 25% of all students were female and that only 20% of the students were full time across all offered degrees.

Cohoon et al. [2] research used qualitative methods to examine why women are not present in greater numbers within United States computing doctoral degree programs. Their research characterizes gender differences.

A study of doctoral dissertations by Lian et al. [8] in a Chinese University studied effects of perceived dissertation quality based on student ages, study periods, modes of study and subject categories, and cross-specialty. The research claimed they found that their categories impacted the quality of doctoral dissertations.

III. DOCTORAL PROGRAM DEMOGRAPHICS

Doctoral Programs targeting working IT professionals have unique strengths and challenges. This section examines recent published demographic data that has been reported by the following: the United States Department of Education, National Science Foundation, the New York State Department of Education, or online by individual universities.

A. Unique IT Professional Student Strengths

Working students offer different perspectives to their education. These students bring the reality into the classroom. As noted by Grossman et al. [11], the research and proposals given by IT professionals are grounded closely related, if not directly integrated, into their full time jobs.

B. Unique IT Professional Student Challenges

Working students have different challenges than traditional full time doctoral degree students. As discussed by Grossman et al. [11], full time working IT professional students are not only required to put their hours into direct IT jobs, they are required to put in additional hours into their studies and research.

C. Doctoral Program Cost

Post undergraduate degrees help to build great achievements in "world's scholars, educators, innovators, and leaders [13]." It is now becoming norm that students are getting post-graduate degrees. Undergraduate degrees are no longer enough for entry level jobs; thus, students are pursuing professional degrees. "In fact, it is projected that about 2.5 million jobs will require either a master's, doctoral, or professional degree between 2008 and 2018 in the United States [13]." Yet, due to lack of financial resources about 40% of all students looking to achieve a doctorate degree do not finish.

Students were surveyed in 2006 by Mendoza et al. [13] to find out how they pay for their doctoral degrees. The findings from the study showed a breakdown of students (partial)

funding sources: 44% through research assistantships, 60% through teaching assistantships, and 48% through fellowships. The survey [13] also indicated that federal loans accounted for 69% of financial aid given to graduate students in 2010–2011. The study showed that the average loan received by students was $16,423 per year.

Without the right amount of financial support students tend not to have continuous enrollment, which causes a normal 3 or 5 year degree to take more time and ultimately decreases the student's chance of completion [13].

D. Part-time Doctoral Programs

Full-time PhD degrees are traditionally multi-year programs where students do not have the opportunity to work full-time jobs. "When Ph.D. programs restrict their enrollments to a small number of full-time students[, they] exclude working professions [14]." There was a study done on graduate students from 1967 until 2000. According to statistics collected by the Department of Education, part-time students represented about 55% of the total graduate-student population through the 1960s and 1970s. As of 2010, it is still relatively large as part-time students amounted to 44% of the total enrolled in graduate programs [14].

Part-time programs, target senior-level managers who are looking for leadership roles in their industry and offer them a program that can potentially help them achieve those roles. Higher education institutions that offer the part-time programs look for people who are thinking about the next stage in their careers. These programs try to support working professionals with flexibility and accommodations to complete their postgraduate degrees. Graduates with a doctorate degree can easily expect to earn six-figure salaries. "A recent salary survey completed by the University of Pittsburgh showed that doctoral candidates received average offers of $100,448 in research schools and $86,769 in teaching schools [15]." The idea for the completion of the part time doctorate degree gives professionals with work experience in their field an opportunity to increase their wages [15].

Part-time doctoral students must have long-term academic readiness which requires the student to have resilience to complete the program. "Because of the fractured student identity of the part-time doctoral candidate, who is usually balancing a range of work, study and family commitments [16];" thus, they must be proactive, have a good support system and have a good plan that will vary based on each student's individual circumstances in order to be successful in completing the program.

E. Doctoral Program statistics across United States

Pace University programs are registered with the New York State Education Department (NYSED). Within NYSED, the Office of Higher Education, approves every curriculum (program) leading to a degree that is registered in the state. Upon reviewing the NYSED Inventory of Registered Programs [17] in Subsection 7, Computer and Information Sciences, there are only two doctoral degree programs awarding a Doctor of Professional Studies. Pace University is registered with a D.P.S. in Computing Studies and is considered to be one of the first D.P.S. programs in Computing in the nation [18]. Syracuse University is registered with a D.P.S. in Information Management. In addition, at the doctoral level in Computer and Information Sciences, there are 18 programs registered as Doctor of Philosophy and one program registered as Doctor of Engineering Science in New York State [17].

The United States Department of Education publishes statistics on educational programs across the United States. The National Center for Education Statistics' Digest of Education Statistics: 2015 [22] provides graduate enrollment in research-based programs by discipline division for selected years from 2007 through fall 2015. As seen in Table 1, during those eight years, the enrollment in graduate computer sciences has increased by 78.65% [19]. As seen in Table 2, the Digest of Education Statistics also provides a range of data on Doctor's degrees conferred by postsecondary institutions, by race/ethnicity and sex of student for selected years, 1976-77 through 2014-15 [20]. The data shows that the number of males earning the degree declined from 1976-77 at 71,709 conferred through 1990-91 at 64,242 conferred. The lowest number of Doctor's degrees conferred to males was in 2002-03 at 62,730 and steadily increased to the highest number in 2013-2014 at 85,585 degrees conferred to males. The number of degrees conferred to females steadily increased every year from 1976-77 with a maximum in 2014-2015 at 93,626 doctor's degrees conferred to females. Females began surpassing males in the total Doctor's degrees conferred in 2011-2012 in the United States.

While females are earning the majority of doctorates in 2015, the National Science Foundation reports in The Survey of Earned Doctorates an exception when it comes to physical and earth sciences, mathematics and computer sciences, and engineering [21].

TABLE 1. UNITED STATES GRADUATE ENROLLMENT IN RESEARCH-BASED DISCIPLINES [19]

Discipline Division			
	2007	***2011***	***2015***
Computer Science	48,246	51,234	86,192

TABLE 2. UNITED STATES DOCTOR'S DEGREES CONFERRED BY POSTSECONDARY INSTITUTIONS [20]

Number of Degrees Conferred to US citizens and nonresident aliens		
Year	***Male***	***Female***
1976-77	71,709	19,509
2000-2001	64,171	55,414
2005-2006	68,912	69,144
2007-2008	73,453	75,925

2011-2012	82,670	87,547
2013-2014	85,585	92,002
2014-2015	84,921	93,626

Section II and Section III above provided insights into our first research question on comparing the Pace University D.P.S. program with other national and international doctoral degree programs.

IV. Pace Dissertation Categorization

The IT working professionals producing dissertations at Pace University bring unique backgrounds and insights into their doctoral research. They all have extensive experience in computing which adds to the speciality of the dissertations. This section analyzes our second research question on what students in the Pace D.P.S. program are researching for their dissertation topics.

A. Methodology

We extracted the abstracts for the dissertations and categorized their text to determine categories of dissertations produced at Pace University. After working to categorize the dissertations by hand, we found that our opinions on categories were individual to the reviewer and that standardization was difficult. As such, we employed machine learning and natural language processing for categorization.

We used TF-IDF processing on the documents and then subjected them to K-means clustering for values of K between 2 and 25.

TF-IDF for "term frequency - inverse document frequency" is a characterization tool for text documents. Each abstract is regarded as a 'bag of words,' as if the meaning of each abstract were implicit in the words used in that abstract and the order of those words were unimportant. Each individual word is them deemphasized according to how often it occurs in the collected abstracts overall.

K-means clustering is an iterative process in which a number K of initial cluster centers is declared, and each datum in a space is then assigned to its nearest cluster. The heterogeneity of each cluster is measured, and the mean point of each of those clusters is declared to be a new cluster center. This process repeats until the heterogeneity fails to decrease or a predetermined maximum number of iterations has been performed.

We applied K-means clustering to the TF-IDF characterization of the abstracts after Zhan, et al.[23]. We dispersed the initial means seven times for each run and captured the lowest heterogeneity score for each value of K as shown in Figure 1 below.

Categorizing documents in this way gives you more well-defined categories as the value of K increases, but the more categories there are, the more difficult reasoning with the model becomes. To balance these two demands, we look for a point on the chart where heterogeneity starts to drop off less quickly with the number of clusters. This appears to happen at around K=6 in Figure 1.

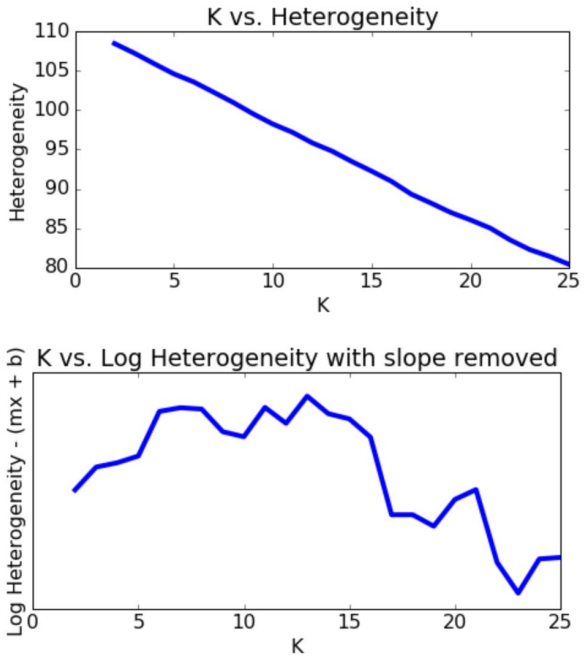

Fig. 1. TF-IDF feature dispersion with the number of clusters

Extracting the top keywords for the lowest heterogeneity run in the 6-cluster case gives us the clusters in Table 3.

While the top plot in Figure 1 looks linear, heterogeneity is simply the sum of the squares of the distances from the cluster center of the cluster members. It will therefore go to zero as K approaches the number of data and each datum gets its own cluster, so is better modeled as an exponential function. We take the log of heterogeneity and subtract the linear slope of that trend to get the bottom plot.

This value -- the difference of log heterogeneity from the best fit line of K versus same -- increases sharply to K=6 then plateaus until K=15 before falling. We therefore select six clusters to best balance interoperability with meaning.

B. Interpretation

These six categories represent six different areas of investigation for D.P.S. dissertations. They align with overall program strategies for the following reasons.

Cluster 0: Software Development
Software development is the professional occupation of many D.P.S. students, and we would expect many of them to do research in this area, especially as they have unusual access for academic work to a operating software development enterprise, which is likely contributing to their tuition.

Cluster 1: Network Security and Data Security Compliance

Again, this is driven by the needs of employers of mid-career technical professionals. Technologists who are not in software development tend to spend a fair amount of time reflecting on security and trying to meet compliance mandates.

Cluster 2: Education
Education makes sense as an area of interest for D.P.S. students because of the prevalence of adjunct and junior faculty in the student body. Many students see a doctorate as a way to advance their careers, however for academics, this is often a concrete requirement to stay in their current job or advance in their career.

TABLE 3. KEYWORDS FOR 6-CLUSTER TF-IDF CATEGORIZATION

#	*Count*	*Top keywords with weight*	*Category*
0	21	software:0.069 agile:0.045 development:0.043 usability:0.039 outsourcing:0.037	Software Development
1	20	security:0.100 cloud:0.054 compliance:0.053 knowledge:0.051 privacy:0.035	Network Security and Data Security Compliance
2	19	students:0.072 science:0.057 computer:0.057 programming:0.055 erp:0.043	Education
3	17	cloud:0.054 genetic:0.053 algorithm:0.043 problems:0.041 algorithms:0.041	Algorithms
4	8	virtual:0.078 projects:0.076 estimating:0.072 cics:0.070 shape:0.068	Project Management
5	29	keystroke:0.054 system:0.041 classification:0.039 biometric:0.037 xml:0.034	Keystroke Biometrics and Text Classification

Cluster 3: Algorithms
Algorithms of course are at the core of Computer Science. One would expect dissertations in this area in a Computing Studies program.

Cluster 4: Project Management
Project manager and technology executives can be expected to tend to focus on these areas, looking for improvements in their processes and platforms.

Cluster 5: Keystroke Biometrics and Text Classification
Keystroke biometrics is a specialty of Pace Seidenberg in particular. It accords with our intuition that one cluster at least would be driven by the interest of the hosting school, which is what we see here.

II. FUTURE WORK

As the popularity of a proven doctorate program for full time working individuals increase, there is a need to perpetually analyze and provide updated statistics to prove or disprove the effectiveness of programs. The first area of future work would be to examine the data using a different statistical methodology or analysis. A second area would focus on student success metrics such as job levels (Jr., Sr., Manager, Executive, etc.), first year annual salaries, and student loans accrued after graduation. A third area would focus on the completion statistics as well as a survey to past cohort members to gain insights into the program experience. The final area of focus would be on furthering the categorization of dissertations, perhaps over time, as this areas has not been widely reviewed.

III. CONCLUSION

The Doctor of Professional Studies (D.P.S.) in Computing is a unique degree program designed for active IT professionals with at least five years of fulltime experience in the computing field. This paper used machine learning and natural language processing to analyze the first 114 dissertations that have been defended at Pace University in the Seidenberg School of Computer Science and Information Systems. The program was officially approved in 1999, with its first dissertation successfully being defended in 2002. We analyzed the topics defended by students to gain an understanding of the topic categories and topic spread produced by the program.

ACKNOWLEDGMENTS

We would like to thank the excellent and dedicated faculty at Pace University that supported the writing of this paper. Specifically, we would like to thank Dr. Fred Grossman for starting this unique program, and to the existing Program Chair, Dr. Charles Tappert, and Associate Program Chair, Dr. Ron Frank, for improving upon an already successful doctorate program for full time professionals. We would also like to thank Michelle Lang for her time helping us from the library in collecting the dissertation abstracts.

REFERENCES

[1] Andrew Bernat and Eric Grimson. 2011. Doctoral program rankings for U.S. computing programs: the national research council strikes out. Commun. ACM 54, 12 (December 2011), 41-43. DOI: https://doi.org/10.1145/2043174.2043203

[2] J. McGrath Cohoon, Zhen Wu, and Jie Chao. 2009. Sexism: toxic to women's persistence in CSE doctoral programs. In Proceedings of the 40th ACM technical symposium on Computer science education (SIGCSE '09). ACM, New York, NY, USA, 158-162. DOI=http://dx.doi.org/10.1145/1508865.1508924

[3] Francisco J. García-Peñalvo. 2013. Education in knowledge society: a new PhD programme approach. In Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality (TEEM '13), Francisco José García-Peñalvo (Ed.). ACM, New York, NY, USA, 575-577. DOI=http://dx.doi.org/10.1145/2536536.2536624

[4] Sabri Serkan Gulluoglu. 2012. Gendered differentiated communication & educational characteristics in online technologies at undergraduate, MA or doctorate degrees. In Proceedings of the 6th Euro American Conference on Telematics and Information Systems (EATIS '12),

Rogerio Patricio Chagas do Nascimento (Ed.). ACM, New York, NY, USA, 435-438. DOI=http://dx.doi.org/10.1145/2261605.2261678

[5] Ilona Heldal, Eva Söderström, Lars Bråthe, and Robert Murby. 2014. Supporting communication within industrial doctoral projects: the thesis steering model. In Proceedings of the 2014 conference on Innovation & technology in computer science education (ITiCSE '14). ACM, New York, NY, USA, 325-325. DOI: http://dx.doi.org/10.1145/2591708.2602680

[6] Kenneth Hoganson and Donald Amoroso. 2011. Inter-disciplinary doctor of science with computer science and information systems. In Proceedings of the 49th Annual Southeast Regional Conference (ACM-SE '11). ACM, New York, NY, USA, 310-311. DOI=http://dx.doi.org/10.1145/2016039.2016122

[7] Donald Joyce and Alison Young. 2004. Developing and implementing a professional doctorate in computing. In Proceedings of the Sixth Australasian Conference on Computing Education - Volume 30 (ACE '04), Raymond Lister and Alison Young (Eds.), Vol. 30. Australian Computer Society, Inc., Darlinghurst, Australia, Australia, 145-149.

[8] Yang Lian, Jingying Chen, and Chunyan Su. 2017. Study on Influencing Factors and Prediction Based on Random Forest of Doctoral Dissertations. In Proceedings of the International Conference on Digital Technology in Education (ICDTE '17). ACM, New York, NY, USA, 64-68. DOI: https://doi.org/10.1145/3134847.3134855

[9] Maria del Carmen Calatrava Moreno. 2012. A qualitative framework for comparison and evaluation of computer science doctoral programs. In Proceedings of the 17th ACM annual conference on Innovation and technology in computer science education (ITiCSE '12). ACM, New York, NY, USA, 398-398. DOI: https://doi.org/10.1145/2325296.2325414

[10] Susan M. Merritt, Allen Stix, Judith E. Sullivan, Fred Grossman, Charles C. Tappert, and David A. Sachs. 2004. Developing a professional doctorate in computing: a fifth-year assessment. In Working group reports from ITiCSE on Innovation and technology in computer science education (ITiCSE-WGR '04). ACM, New York, NY, USA, 42-46. DOI: http://dx.doi.org/10.1145/1044550.1041654

[11] Fred Grossman, Charles Tappert, Joe Bergin, and Susan M. Merritt. 2011. A research doctorate for computing professionals. *Commun. ACM* 54, 4 April, 133-141. DOI: https://doi.org/10.1145/1924421.1924450

[12] Cagle, Tammie. 2017. “TOP 50 DOCTORATE IN BUSINESS MANAGEMENT PROGRAMS.” Top Management Degrees, Retrieved from:www.topmanagementdegrees.com/rankings/best-doctorate-business-management/ on 20 Nov. 2017.

[13] Mendoza, Pilar, et al. 2014. "Within-Year Retention among Ph.D. Students: The Effect of Debt, Assistantships, and Fellowships." Research in Higher Education, vol. 55, no. 7, 01 Nov., pp. 650-685.

[14] CASSUTO, LEONARD. 2013. "The Part-Time Phd. Student." Chronicle of Higher Education, vol. 60, no. 6, 11 Oct., pp. A26-A27.

[15] Patricia V., Rivera. 2017. "Part-Time Doctoral Program in Dallas Area Targets Working Professionals." Dallas Morning News, the (TX), n.d.

[16] Watts, Jacqueline H. 2008. Challenges of Supervising Part-Time PHD Students: Towards Student-Centred Practice. Teaching in Higher Education, vol. 13, no. 3, pp. 369-373. EBSCOhost, doi:10.1080/13562510802045402.

[17] New York State Education Department. 2017. Retrieved from: http://www.nysed.gov/coms/rp090/IRPSL1

[18] Wikipedia. 2017. Pace University D.P.S. Program. Retrieved from: https://en.wikipedia.org/wiki/Doctor_of_Professional_Studies

[19] National Center for Education Statistics. 2017. Digest of Education Statistics 2015. Retrieved from: https://nces.ed.gov/programs/digest/d15/tables/dt15_311.90.asp

[20] National Center for Education Statistics. 2017. Digest of Education Statistics 2015. Retrieved from https://nces.ed.gov/programs/digest/d16/tables/dt16_324.20.asp?current=yes

[21] National Science Foundation. 2017. The Survey of Earned Doctorates. Retrieved from: https://www.nsf.gov/statistics/2017/nsf17306/report/about-this-report.cfm

[22] National Science Foundation. 2017. The Survey of Earned Doctorates. Retrieved from: https://nces.ed.gov/programs/digest/d15/index.asp

[23] John B. Horrigan. 2016. Lifelong Learning and Technology. Pew Research Center Report. Retrieved from: http://www.pewinternet.org/2016/03/22/lifelong-learning-and-technology