

Texas A&M University-San Antonio

Digital Commons @ Texas A&M University-San Antonio

Computer Information Systems Faculty
Publications

College of Business

2023

Mining Health Informatics Job Advertisements: Insights for Higher Education Programs and Job Seekers

Ahmed El Noshokaty

Mohammad A. Al-Ramahi

Omar El-Gayar

Abdullah Wahbeh

Tareq Nasrallah

Follow this and additional works at: https://digitalcommons.tamusa.edu/cis_faculty



Part of the [Computer Sciences Commons](#)

Mining Health Informatics Job Advertisements: Insights for Higher Education Programs and Job Seekers

Ahmed Elnoshokaty
California State University, San Bernardino
ahmed.elnoshokaty@csusb.edu

Mohammad Al-Ramahi
Texas A&M University-San Antonio
mohammad.abdel@tamusa.edu

Omar El-Gayar
Dakota State University
omar.el-gayar@dsu.edu

Abdullah Wahbeh
Slippery Rock University
abdullah.wahbeh@sru.edu

Tareq Nasrallah
Northeastern University
t.nasrallah@northeastern.edu

Abstract

This paper used web scraping and data mining to analyze 831 health informatics job advertisements on indeed.com. Results showed that 87% of jobs explicitly required a college degree in a related field, 41% of jobs preferred a graduate degree, while 29% preferred or required professional certification. The analysis showed that preferred skills were analytics problem solving, communication skills, oral communication, interpersonal skills, project management, statistics, and critical thinking. The analysis also showed that college degrees, certifications, and the above-mentioned skill set are in high demand for working in the field of health informatics, especially in states with large populations and strong economies. Our results inform curriculum development of health informatics programs in higher education, which helps map knowledge units across the curricula to bridge the skills gap and meet employers' expectations. At the same time, the results help job seekers familiarize themselves with what employers seek in a successful candidate.

Keywords: Health Informatics, Text Mining, Job Posts, Job Seekers, Employers

1. Introduction

The Bureau of Labor Statistics projects that jobs in health care will grow before 2026 by about 125 thousand openings, with middle skilled jobs representing about 69 thousand of them (L'Heureux, 2019). According to the US Institute of Medicine (IOM), poor use of data and data integration problems are among the reasons for the \$750 billion (out of \$2.5 trillion expended) annual excess cost in healthcare (Smith et al., 2013).

More resources were devoted to those programs to improve data management in health care, including \$2 billion for hospitals and healthcare providers to adopt and use health informatics and electronic health records (Blumenthal, 2009; Buntin et al., 2010).

Health informatics is a field that involves working with data as well as data analytics tools to enhance the quality of health and healthcare. However, despite the high demand for health informatics in the current century, the health informatics field is not routinely taught to health professionals (Fridsma, 2018).

Health informatics skills fill the knowledge gap of information technology and information management jobs seekers. Health informatics professionals are required to improve their skills to meet the needs of health informatics companies and government agencies (Hasman & Albert, 1997).

Existing research (Ashrafi et al., 2014; Fridsma, 2018) analyzed health informatics university programs and course curricula to better prepare public health professionals with the knowledge and skills needed. Actions for academia were suggested to advance the mission of public health companies and providers.

However, most existing studies neglect job posting and mainly focused on surveying health informatics professionals or analyzing university health informatics curricula (Ashrafi et al., 2014; Garde et al., 2006; Hasman & Albert, 1997). The few studies that utilized job posting (Joshi et al., 2021) were limited in terms of the methodology followed, where the analysis was completed manually using content analysis.

Mining job advertisements is crucial to understand the current needs for healthcare informatics from the employer's perspective. It also could help enhance curricula development (Singh Dubey & Tiwari, 2020) by having better understanding of the subject matter and skills required in the health informatics field. Therefore, ensuring that graduates have the correct set of

knowledge, skills, and mastery of tools will pave the way to a successful career and positively impact the healthcare domain (Ogungbemi et al., 2021).

Accordingly, this research aims to address the following question: What are employers' expectations from health informatics job seekers? More specifically, we aim to identify important degrees, certifications, and skills expected by employers from prospective job seekers. This is accomplished through mining health informatics jobs on popular online job search engines such as indeed.com.

The findings from this study will help inform the design and continuous improvement effort of health informatics programs and professional training initiatives in a manner that better aligns with employers' expectations. In addition, the findings can assist health informatics job seekers in identifying the knowledge, skill sets, and degree expectations for successful careers in health informatics.

The remainder of this paper is organized as follows. The background and related works section surveys relevant literature related to the analysis of skills, certificates, and degrees required by the job market. The methodology section details the steps followed for data collection and analysis. The results section presents findings, then we conclude the paper with discussion and summary of findings and limitations.

2. Background and Related Works

Information technology has transformed the healthcare industry and changed the way medical professionals serve populations. Similarly, healthcare informatics has significantly helped improve the quality of population health and the healthcare sector. However, despite the significance of these fields, they are not usually taught to be health professionals (Fridsma, 2018).

A limited number of studies have attempted to characterize the field of healthcare informatics and provide a big picture of the knowledge, skills, and education required to support the increasing needs of healthcare organizations. For example, Joshi et al., (2021) analyzed health informatics related job postings using content analysis of jobs posted on Indeed.com to characterize the knowledge, skills, and expertise required by the employers of the informatics workforce. A total of 321 job postings were analyzed and results showed that analyst was the most common job category, followed by director and faculty/researcher. Furthermore, the most common hiring agencies were health systems and hospitals, followed by academic institutes. In addition, 12% of the jobs did not list the degree, and 19% indicated masters as the preferred degree. In addition, the results showed that many

postings required skills related to “requirement analysis, business processes and working in multidisciplinary teams' skills.”

Hess (2021) explored the skills needed by health informatics job seekers and how graduate degree competencies align with employers' perception of these skills. Following a conceptual framework, the author collected data using interviews with executives, practitioners, and hiring managers in the field of health informatics. Qualitative data analysis showed that a graduate degree increases employability. The author also recommended a “student-to-employee pipeline” to help grow the workplace in the field of health informatics.

Butler-Henderson et al., (2020) analyzed the demographic, educational, and occupational characteristics of health information technology specialists in Australia and New Zealand by surveying a total of 227 participants. Statistical analysis showed that 48.5% of the respondents have a college degree, with 22.7% of the respondents in information technology followed by 20.2% in computer science and 19.3% in health information management. About 30% of the respondents' role titles were classified as leadership roles, while the remaining roles were related to information/computer technology, systems, or science functions. About 82% of the respondents selected health information technology specialists as an occupational group.

Ashrafi et al., (2014) analyzed the differences between health informatics and big data in the field of healthcare by analyzing curricula in 127 higher education institutions. Results showed that only 16 universities offer an undergraduate degree in health informatics, 72 offer master's programs, and 13 offer Ph.D. programs. Only 12 universities offer undergraduate certificates, while 48 offer graduate certificates. Overall, 313 courses focus on technology, 66 emphasize clinical knowledge, and 390 courses focus on other aspects of health informatics. Most of the programs offer one to three courses that are related to health informatics, which technically falls short in terms of providing the necessary knowledge in the field.

Few countries have policies to integrate informatics in education (Otero et al., 2020) despite the fact that interdisciplinary programs are believed to have great academic value (Ogungbemi et al., 2021). It is widely recognized that healthcare information technology and healthcare informatics professionals should have interdisciplinary knowledge related to “information technology, healthcare, business, management, and other disciplines” (Hersh, 2010). Studies showed that a gap exists in the “competencies of public health graduates in meeting the requisite skills and demands of the dynamic public health workforce” (Joshi et al.,

2021). Job seekers must learn a set of skills and competencies that can help them succeed in health informatics jobs (Hess, 2021). Unfortunately, current health informatics job seekers' assessments have shown that they “lack adequate health informatics skills” (Joshi et al., 2021). As a result, the healthcare informatics curriculum would benefit more when they are designed around training, tools, and knowledge that can help enhance the domain of healthcare (Ogungbemi et al., 2021).

Overall, studies that addressed skills and knowledge needed in the health informatics domain are limited in terms of the approaches followed for data collection and analysis, where most studies were based on survey data and did not explore the potential of large-scale data available on job postings websites. Even the few studies that utilized job postings website data, were limited in terms of the analysis process, where job postings were manually analyzed using content analysis, which is considered a tedious and time-consuming process thereby limiting the scope of the research.

3. Methodology

A web crawler is a bot that collects publicly available data on the internet. We developed a crawler to collect health informatics related posts from indeed.com, a well-known website for jobs and internships. Indeed was selected as it is the top employment Website visited by 25 million unique job seekers every month and had options for job seekers to express their opinions about the companies they worked for (Sainju et al., 2021).

The web crawler retrieved health informatics jobs at indeed.com using health informatics jobs related keywords “health informatics” or “clinical informatics” (Hersh, 2010). We collected all the details about each job post, including the job title, job location, job posting date, job description, institution name (company or government agency posting the job), and indeed institution ratings (average anonymous ratings of institution’s employees on indeed.com).

As shown in figure 1, the collected posts’ descriptions were preprocessed and prepared for analysis. First, all whitespace (such as new lines and extra spaces), punctuation, and HTML tags were removed. Second, the text was converted to lower case, and stop words were removed. Stop words are basically a set of commonly used words in any language like “the,” “into,” “just,” and “keep.” By removing words that are very commonly used in each language, we could focus only on the important words instead and improve the accuracy of the text processing. Next, we applied

lemmatizations for all words to reduce inflectional word forms to linguistically valid lemmas.

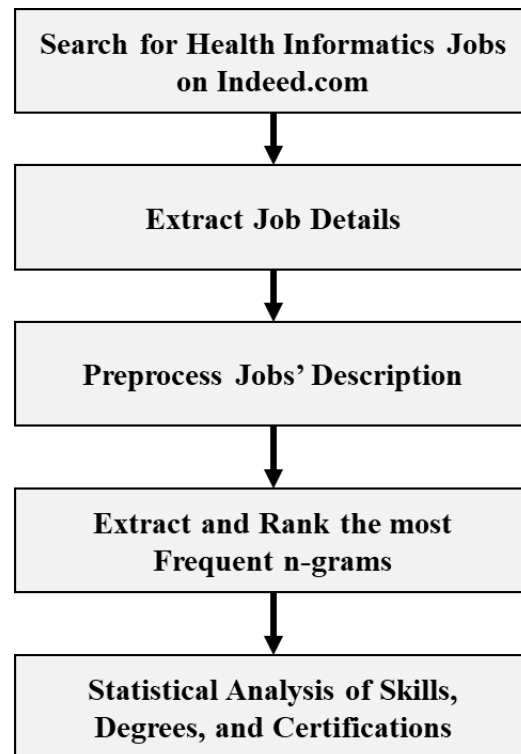


Figure 1. Data collection and mining process

To extract undergraduate or graduate degree requirements, we searched for keywords like “bachelor's degree,” “master's degree,” “BSc,” “MSC,” “MBA,” “MD,” “doctorate,” and “PhD.” and related variations. To extract professional certification requirements, we looked up unigram, bi-grams, tri-grams, four-grams, five-grams, and six-grams in the sentence where the keywords “certified,” “certification,” or “certificate” was listed and determined the highest frequency of term occurrence to detect the certifications with the highest demand in the job market such as the certified professional coder (CPC), project management professional (PMP), and registered health information technician (RHIT).

To assess skills requirements, we looked for unigram, bi-grams, tri-grams, four-grams, five-grams, and six-grams in job keywords and job descriptions and determined the highest frequency of term occurrence to detect the skills with the highest demand in the job market like “Excel,” “analytical problem solving & troubleshooting,” “communication skills,” “oral communication,” “interpersonal skills,” “statistics,” “structured query language (SQL),” “SAP,” and “critical thinking.”

4. Results

The crawler collected 831 jobs from 592 institutions that were available on indeed.com as of May 23rd, 2022 (Table 1). As shown in table 1, 87% of the health informatics jobs required a university or college degree, 41% of the jobs posted required or preferred graduate degree, while 29% of the jobs posted required or preferred some certifications.

Table 1: Degree and certifications in health informatics job postings

Degree	Postings
University or college degree required	87% (726)
Prefer or require a graduate degree	41% (343)
Prefer or require certification	29% (242)

Table 2 shows the most popular certifications required by the employer. Healthcare related certification such as EHR and EPIC, RHIA, and RHIT were listed as required or preferred certificates by a total of 102 jobs. Project management professional certification was listed in a total of 18 jobs, while Lean Six Sigma certification was required or preferred by 7 jobs.

The remaining jobs that required or preferred certification (account for 14% of the jobs) were generic and did not explicitly specify the type of certification. Examples of generic certification include "certification relevant to training," "industry designation certifications," "industry-focused certifications," "certification from a national informatics certifying body," and "specialty certification." In addition, some employers stated that they would provide support for obtaining certifications as part of the professional development benefits to attract candidates.

Table 2: Most frequent certification requested in health informatics job postings

Certifications	Postings
EHR including EPIC Certifications (Care and Cogito)	5% (42)
Certified Coder CC and Certified Professional Coder CPC	3% (28)
Project Management Professional PMP	2% (18)
Registered Health Information Administrator RHIA	2% (18)
Registered Health Information Technician RHIT	2% (14)
Lean Six Sigma	1% (7)

Table 3. shows the top required or desired skills and qualifications by employers. Analytical problem skills are reflected in three-fourth (75%) of the job postings, followed by communication and oral skills, project

management skill, statistics, interpersonal skills, and critical thinking with 8% of the job postings.

Table 3: Most frequent skills requested in health informatics job postings

Skills	Postings
Analytical Problem Solving	75% (623)
Communication Skills	41% (337)
Oral Communication	29% (238)
Project Management	24% (199)
Statistics	24% (198)
Interpersonal Skills	21% (174)
Critical Thinking	8% (67)

Table 4 shows a comparative analysis of the required or preferred skills associated with undergraduate and graduate and degrees.

Table 4: Most frequent skills associated with requested degree types in health informatics job postings

Skills	Under-graduate	Graduate
Analytical Problem Solving	47% (392)	30% (255)
Communication Skills	29% (248)	15% (130)
Oral Communication	20% (165)	14% (120)
Project Management	18% (155)	10% (85)
Statistics	17% (141)	14% (117)
Interpersonal Skills	15% (128)	9% (71)
Critical Thinking	6% (49)	3% (26)

Table 5 shows the data management and programming tools that are preferred by employers based on the job descriptions.

Table 5: Most frequent tools expected in health informatics job postings

Tools	Postings
Excel	62% (514)
Structured Query Language (SQL)	18% (150)
Python	9% (74)
Tableau	9% (73)
Java	4% (29)
R	3% (24)
Power BI	3% (13)
SPSS	2% (14)

Most jobs are offered on the west and east coasts as well as in Texas, as shown in Figure 2. A total of 221 (27% of jobs) jobs provided the flexibility of working remotely. Figure 3. shows the states with the highest

percentage of jobs requesting college degree requirements.

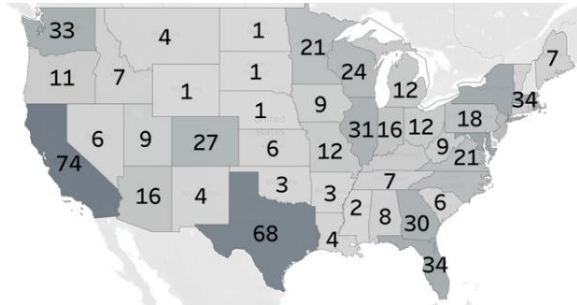


Figure 2. States with a high supply of health informatics jobs

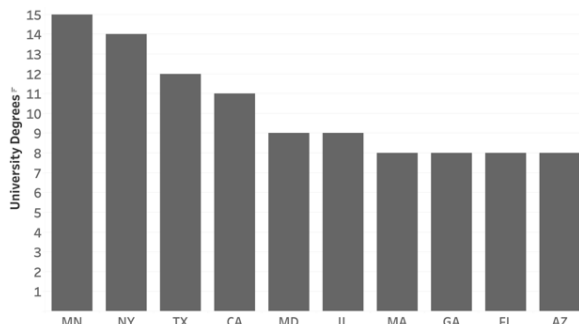


Figure 3. States with highest university degrees demand in jobs

Many jobs across all states also require certification. Figure 4 shows the states with the highest supply of jobs requesting certifications. The figure shows that Texas and California were the highest states with job pots requesting health informatics-related certification.

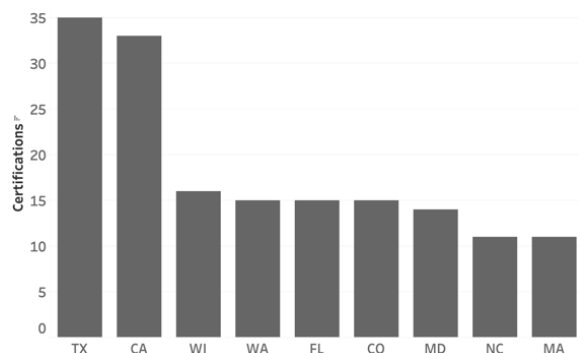


Figure 4. States with highest certifications demand in jobs

5. Discussion

Majority of the job postings required a university degree, whether an undergraduate degree (87% of jobs) or graduate degree (41% of the jobs). Certifications also

have a significant presence (almost one third (29%) of job postings). This is consistent with (Rob & Roy, 2013) for the broader information systems domain where certification is considered a key element that can offer job seekers competitive advantage. Certifications can complement or replace formal education requirements (Hollister et al., 2017). According to Mardis et al., (2018), certification could be used as an evaluation criterion by employers since it can signal the applicant’s aptitude and potential.

The desired certificates focus on administrative and managerial such as project management professional certification, a popular certification that equips professionals with the knowledge needed to play an important role in leading and succeeding with a project (Jinlan et al., 2019). The registered health information administrator (RHIA) and registered health information technician (RHIT) certifications are designed for students who completed an accredited health information management program (DeAlmeida et al., 2019). There is also evidence for technology- (and vendor dependent) certification such as Epic clarity is a “platform that is used separately from the Epic EMR platform so that complex, data-intensive reports can be generated without interfering with the EMR platform” (Edelstein, 2020).

Skills captured in the job postings predominantly emphasize analytical problem solving and soft skills. This could be a reflection of digital transformations happening in the healthcare domain and that inevitably expects having a skilled workforce to support the use of technologies used in healthcare delivery (Butler-Henderson et al., 2020).

In both undergraduate and graduate programs, matching the employers' desired skills should be reflected in the health informatics curricula and extra-curricular activities. The need exists to ensure that health informatics curricula integrate skills and knowledge from different disciplines, including computing, statistics, and project management, in addition to soft skills that can help job seekers succeed in the industry. One way to assure the integration of these skills into curricula could be the adoption of the competencies approach to align offerings with required workforce skills in health informatics (Hess, 2021). Further, partnerships with the industry could afford internship opportunities as well as exposure to ‘real-world’ problems for various student projects.

The most frequent tools identified in job postings include general tools such as Excel, programming languages, statistical tools, and business intelligence (BI). Python is the most referenced programming language desired by the employer, followed by Java, R, and SPSS. Overall, these tools indicate the importance of statistics as reflected in the list of desired skills and

knowledge by employers. In addition, required BI tools such as Tableau and Power BI emphasize the need for data visualization knowledge.

The findings indicate that health informatics programs need to provide a broad exposure to a variety of tools and technologies. These can span from general tools such as Excel, to programming and specialized analytics and visualization tools.

This study is not without any limitations. From a methodological standpoint, the "AS" abbreviation for associate in science is a homonym with the stop word "as", which could not be detected. In addition, we might have missed synonyms of skills due to keyword matching and occurrence count.

In the scope of this paper, we could not study in depth the dynamics of the hiring process nor contribute to identifying who gets hired. A future extension of this work is to survey employers to understand the hiring process. Specifically, studying the relationship between job requirements (such as university degree and certification requirements) and job benefits (such as salary, flexibility, and location). Also, we aim in the extension to learn more about the screening and interviewing process for health informatics jobs and study how candidates are being selected. Future research could also widen the scope and study jobs in other related IT fields to complement our findings since we only focused on health informatics-related jobs in this study.

6. Conclusion

This paper analyzed job postings related to health informatics positions. The results highlight the prevalence of university degrees and recognize the growing importance of industry-specific certifications. The results also highlight the importance of problem solving and analytic skills along with soft skills. Employers also expect exposure to a variety of tools including programming languages, business intelligence, and statistical tools, as well as general tools such as Excel.

It is also evident that the health informatics field is interdisciplinary and requires knowledge of technology, statistics, healthcare, and project management. The findings from this study inform curriculum development for universities offering degrees related to health informatics (curriculum development). Furthermore, these findings provide guidance to job seekers as they prepare for careers in health informatics.

Overall, an important implication of this research is the need for university-industry collaboration in the development and continuous improvement of health informatics curricula to align the health informatics programs in higher education with industry needs.

Integrating coursework that prepares students to acquire relevant industry certification will make students more employable. In the digital and AI era, where nearly all future jobs will be knowledge/technological-intensive, there will be a growing need for T-shaped upskilling, with breadth and depth.

The need exists for higher education to turn to innovative, quick, and scalable solutions to close the skills gap across a variety of curricula. In this study, we introduced the potential of text mining to automatically analyze job posts to bridge this gap in the health informatics field (and other fields in general).

7. References

- Ashrafi, N., Kuilboer, J.-P., Joshi, C., Ran, I., & Pande, P. (2014). *Health Informatics in the Classroom: An Empirical Study to Investigate Higher Education's Response to Healthcare Transformation*. 25, 12.
- Blumenthal, D. (2009). Stimulating the Adoption of Health Information Technology. *New England Journal of Medicine*, 360(15), 1477–1479. <https://doi.org/10.1056/NEJMp0901592>
- Buntin, M. B., Jain, S. H., & Blumenthal, D. (2010). Health Information Technology: Laying The Infrastructure For National Health Reform. *Health Affairs*, 29(6), 1214–1219. <https://doi.org/10.1377/hlthaff.2010.0503>
- Butler-Henderson, K., Gray, K., Day, K., & Grainger, R. (2020). Defining the Health Information Technology discipline: Results from the 2018 Australian and New Zealand censuses. *Proceedings of the Australasian Computer Science Week Multiconference*, 1–6. <https://doi.org/10.1145/3373017.3373043>
- DeAlmeida, D. R., Houser, S. H., Wangia-Anderson, V., Fenton, S. H., Hazelwood, A., Barefield, A. C., Freeman, J. M., Jones, L. M., Bakuzonis, K., & Hamada, D. L. (2019). An Exploratory Study Demonstrating the Health Information Management Profession as a STEM Discipline. *Perspectives in Health Information Management*, 16(Summer), 1a.
- Edelstein, J. M. (2020). *The Impact of Occupational Therapy Acute Care Services on Readmission Rates for Patients in Medicare's Hospital Readmission Reduction Program*. Virginia Commonwealth University.
- Fridsma, D. B. (2018). Health informatics: A required skill for 21st century clinicians. *BMJ*, k3043. <https://doi.org/10.1136/bmj.k3043>
- Garde, S., Harrison, D., Huque, M., & Hovenga, E. J. (2006). Building health informatics skills for health professionals: Results from the Australian Health Informatics Skill Needs Survey. *Australian Health Review*, 30(1), 34. <https://doi.org/10.1071/AH060034>
- Hasman, A., & Albert, A. (1997). Education and training in health informatics: Guidelines for European curricula. *International Journal of Medical Informatics*, 45(1–2), 91–110. [https://doi.org/10.1016/S1386-5056\(97\)00039-7](https://doi.org/10.1016/S1386-5056(97)00039-7)
- Hersh, W. (2010). The Health Information Technology Workforce: Estimations of Demands and a Framework

- for Requirements. *Applied Clinical Informatics*, 01(02), 197–212. <https://doi.org/10.4338/ACI-2009-11-R-0011>
- Hess, P. C. (2021). *A Healthcare Informatics Skills Framework for Employers, Educators, and Practitioners: A Qualitative Approach - ProQuest*. https://www.proquest.com/openview/483d7e17d1874dff3438ae0310b9fa8e/1?casa_token=1PzAWZ37JYIAAA:AA:qhp6KdI1M42el6tLVMOTCki-em9JR0ah4KtIIRiX1sGL97bAcI-0YlgOtnDI4-uwADe3jzNE3aE&cbl=18750&diss=y&pq-origsite=gscholar&parentSessionId=gAYrdj%2BBOx%2B1OlgWmq2%2Bz77uQ6nMRCCID6el4JMQPE%3D
- Hollister, J. M., Spears, L. I., Mardis, M. A., Lee, J., McClure, C. R., & Liebman, E. (2017). Employers' perspectives on new information technology technicians' employability in North Florida. *Education + Training*, 59(9), 929–945. <https://doi.org/10.1108/ET-02-2017-0019>
- Jinlan, Z., Jing, H., Dengsheng, W., & Jianping, L. (2019). Mechanism of How Project Management Professional Certification Influences Employees' Competency and an Empirical Analysis in Support. *Management Review*, 31(4), 138.
- Joshi, A., Bruce, I., Amadi, C., & Amatya, J. (2021). Developing Evidence-based Population Health Informatics curriculum: Integrating competency based model and job analysis. *Online Journal of Public Health Informatics*, 13(1), e10. <https://doi.org/10.5210/ojphi.v13i1.11517>
- L'Heureux, T. J. (2019). *Strangers in Their Own Land: The Participation of Arizona's Hispanic Immigrants in Middle-skilled Healthcare Occupations*. University of Chicago.
- Mardis, M. A., Ma, J., Jones, F. R., Ambavarapu, C. R., Kelleher, H. M., Spears, L. I., & McClure, C. R. (2018). Assessing alignment between information technology educational opportunities, professional requirements, and industry demands. *Education and Information Technologies*, 23(4), 1547–1584. <https://doi.org/10.1007/s10639-017-9678-y>
- Ogunbemi, O., Gupta, S., & Sneha, S. (2021). *Assessment of Healthcare Informatics Education in USA: The Current State and Comparisons across Program Types* (p. 16). Kennesaw State University.
- Otero, P., Leikam, M., Gonzalez, Z., Marin, H. de F., Aravena, I. P., & Zawadzki, S. (2020). Informatics Education in Latin America. In E. S. Berner (Ed.), *Informatics Education in Healthcare: Lessons Learned* (pp. 167–182). Springer International Publishing. https://doi.org/10.1007/978-3-030-53813-2_13
- Rob, M., & Roy, A. (2013). The Value of It Certification: Perspectives from Students and It Personnel. *Issues In Information Systems*. https://doi.org/10.48009/1_iis_2013_153-161
- Sainju, B., Hartwell, C., & Edwards, J. (2021). Job satisfaction and employee turnover determinants in Fortune 50 companies: Insights from employee reviews from Indeed.com. *Decision Support Systems*, 148, 113582. <https://doi.org/10.1016/j.dss.2021.113582>
- Singh Dubey, R., & Tiwari, V. (2020). Operationalisation of soft skill attributes and determining the existing gap in novice ICT professionals. *International Journal of Information Management*, 50, 375–386. <https://doi.org/10.1016/j.ijinfomgt.2019.09.006>
- Smith, M., Saunders, R., Stuckhardt, L., & McGinnis, J. M. (2013). Best Care at Lower Cost: The Path to Continuously Learning Health Care in America. *The National Academies Press*, 381.