

Program Review and Assessment Report

Indiana University School of Informatics and Computing, IUPUI

2017–2018

The following degree programs are reviewed in the 2017–2018 academic year:

- Bachelor of Science in Informatics
- Master of Science in Applied Data Science

Schedule

	2015– 2016	2016– 2017	2017– 2018	2018– 2019	2019– 2020	2020– 2021	2021– 2022
Health Information Management BS	✓			✓			✓
Informatics BS			✓			✓	
Media Arts and Science BS	✓			✓			✓
Bioinformatics MS	✓			✓			✓
Health Informatics MS		✓			✓		
Human–Computer Interaction MS		✓			✓		
Applied Data Science MS			✓			✓	
Master of Library Science		✓			✓		
Media Arts and Science MS		✓			✓		

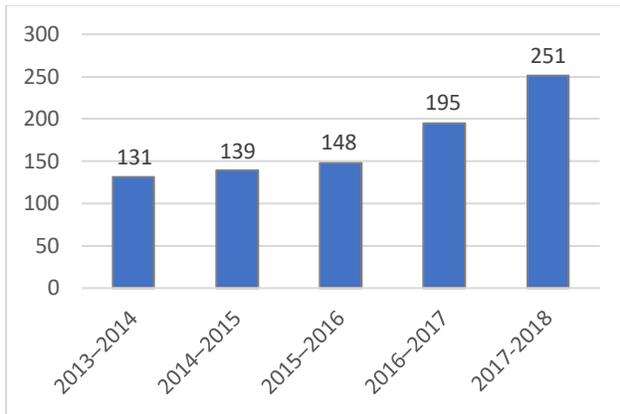
Bachelor of Science in Informatics
Program Review and Assessment Committee Report
Department of Human-Centered Computing
Indiana University School of Informatics and Computing

1. Academics

Given that the 2017–2018 academic year was dedicated to the revision of the program’s learning outcomes, the assessment of the old learning outcomes as a direct measure is not reported. Enrollments, degrees conferred, four-year graduation rate, and one- and two-year graduation rate are provided as indicators of academic success. Starting from the 2018–2019 academic year, the revised learning outcomes will be assessed.

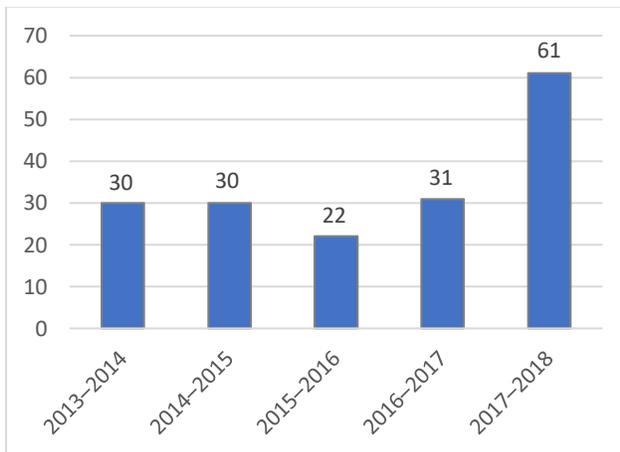
1.1. Enrollment

The Bachelor of Science in Informatics program has experienced an increase in enrollment over the past five academic years from 131 students in fall 2013 to 251 students in fall 2017.



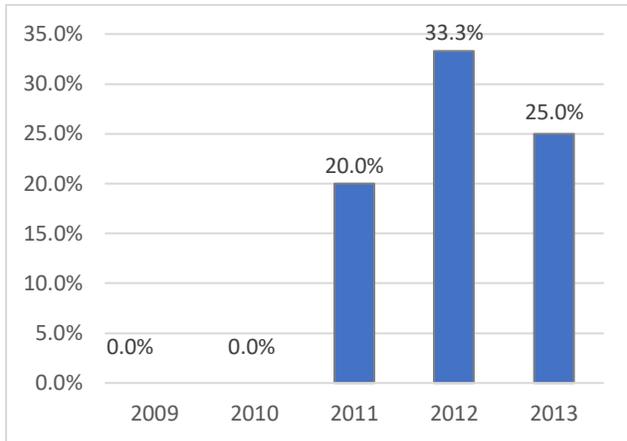
1.2. Degrees Conferred

The number of students graduating with a Bachelor of Science degree in Informatics has more than doubled in the last five years.



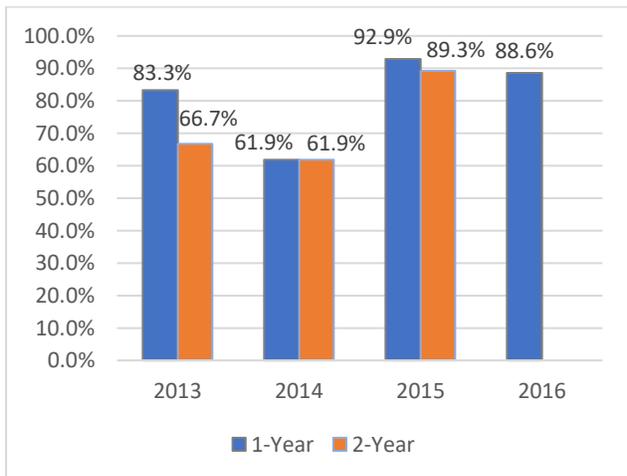
1.3. Four-Year Graduation Rate

The four-year graduation rate has increased from 0% for the 2009 cohort to 25% for the 2013 cohort. The graduation rate for the 2014 cohort is expected to be much higher than 25%.



1.4. One- and Two-Year Retention Rates

Retention rates follow an upward trend. Data from 2012 is excluded because the cohort had only 3 students.

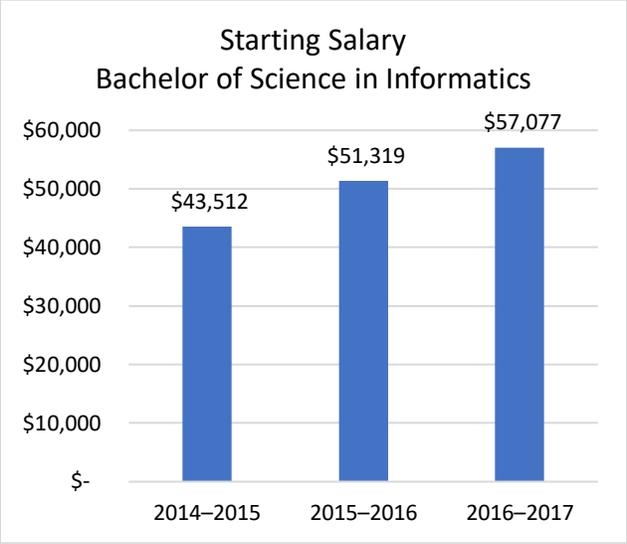


2. Careers

As an indirect measure for program assessment, we consider starting salaries and the number of students placed in major for the past three years.

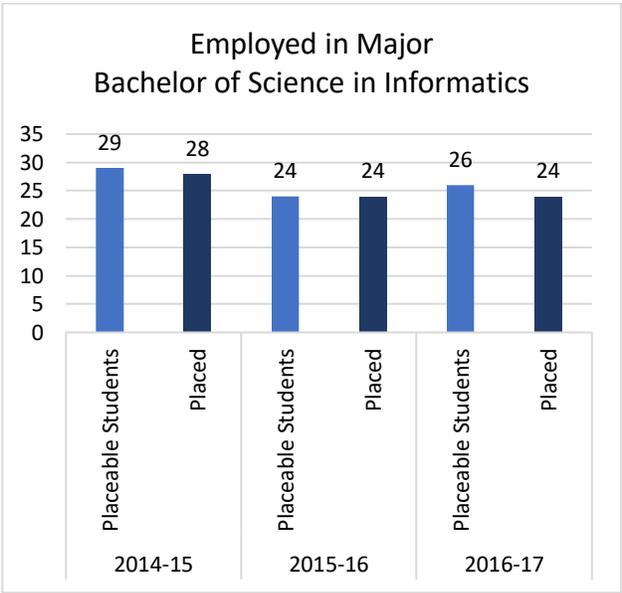
2.1. Starting Salaries

Average starting salaries have increased three years in a row.



2.2. Students Placed in Major

Between 92% and 100% of graduates were placed in a job within their major.



Indiana University School of Informatics and Computing, IUPUI
Department of Human-Centered Computing
PRAC Assessment Project

Name and rank/title of Project Director(s):

Sara Anne Hook, M.B.A., J.D., Professor of Informatics
Liugen (Louie) Zhu, Ph.D., Senior Lecturer, Informatics

Program/Track and School:

Bachelor of Science in Informatics
Indiana University School of Informatics and Computing (SoIC)

Campus Address:

535 West Michigan Street, IT 589
Indianapolis, IN 46202

Phone: (317) 278-7690
Fax: (317) 278-7669
Email: sahook@iupui.edu, louiezhu@iupui.edu

Project Title: Reviewing, Revising and Implementing the Program-Level Outcomes (PLOs) for the B.S. in Informatics

Project Dates: January 1, 2017 – June 30, 2018

Project Summary

The intended outcome of the project was to develop an updated and carefully considered list of program-level outcomes (PLOs) for the B.S. in Informatics that is specifically tailored to this degree and that included the input from faculty members, administrators, career services professionals, and potential employers. The revision of the PLOs is essential at this point in the evolution of SoIC, especially given the feedback from the Program Review team in 2016 (Appendix A: Initial Learning Outcomes for the B.S. in Informatics)

This project combined a literature review, a syllabus review, a survey of faculty members, and interviews with key stakeholders to examine existing program-level outcomes (PLOs) for the B.S. in Informatics and used the results of this inquiry to revise the PLOs, which were inadequate, outdated, and difficult to measure.

The project directors have consolidated their findings from these activities and have prepared a draft proposal for the updated PLOs for the B.S. in Informatics degree that is a mixture of revisions to the previous PLOs (with more action verbs and measurable outcomes) and new PLOs that reflect the reality of current technology and predictions about technology in the future as well as the essential soft and professional skills needed for the current workplace. The intent is that once the PLOs are approved by the faculty members, Department Chair, and other administrators, faculty members will incorporate the revised PLOs into syllabi for all courses that are part of the B.S. in Informatics and ensure that they are matched with appropriate assessment methods to clearly demonstrate that the PLOs are being achieved.

Data Collection for PLO Revision

Literature Review

Reviewed literature on trends in computing and information technology and the learning outcomes of closely related programs, including computer science, information technology, computer technology, data studies, and information science.

Syllabus Review

Gathered and reviewed all syllabi for core courses and selected elective courses in the B.S. in Informatics program to see which syllabi complied with the request to include the PLOs and whether the PLOs were tied to specific assignments and other assessment methods (Appendix B: Syllabus Review).

Survey of Faculty Members

Prepared and distributed a survey to 14 faculty members teaching core and elective courses in the B.S. in Informatics program to gather their thoughts on the existing PLOs and how they should be revised (Appendix C: Survey Questionnaire of Faculty Members). The faculty was given three weeks to respond to the survey. At the end of the survey, four faculty members responded to the survey. The following table shows the summary of the survey results (Appendix C: Survey Questionnaire of Faculty Members).

1. Skills and knowledge to prepare students	Programming Algorithmic thinking UX design Data integration, analytics, transformation, and visualization Tools and practices used in industry Ethical and proper use of technology
2. Short-term and long-term trend in IT fields	Big data Machine learning Augmented reality Mobile apps Data integration and governance Data parsing Cloud computing, data warehousing
3. Technology skills a graduate should have	Scripting and programming Relational and nonrelational databases UX design Data analysis Hands-on business intelligence, extract-transform-load (ETL), and analytic tools Data integration and mining tools
4. Soft skills a graduate should have	Communication Teamwork strategies Design ethics Organization skill Ability to teach and train others Project management
5. Skills and experience that have helped a graduate in finding a position	Technology knowledge Web design and development Data structures Data indexing, mining, and visualization Critical thinking
6. Initial PLOs to be retained	Measurable project-based learning activities Case studies for assessment

Interviews with Key Stakeholders

Interviewed the Department Chair, two of SoIC’s career services staff members, and two industry professionals who have hired graduates of our program about what the marketplace needs in the way of properly prepared, work-ready graduates. The questions were the same as those used to survey faculty members. The interviews were conducted in person or by phone. Responses were similar to those we received through the faculty survey, with more emphasis on soft skills and less on technical skills. The following table shows a summary of the interview responses.

<p>1. Skills and knowledge to prepare students</p>	<p>Project management Business aspects of technical knowledge Some SQL Internships Service learning projects Case studies/real-life scenarios Presentation skills I210/I211 courses Adapt technology to different settings Privacy Cybersecurity Business-to-business (B2B) IT Business-to-customer (B2C) IT Different types of IT (e.g., interactive voice response systems)</p>
<p>2. Short-term and long-term trend in IT fields</p>	<p>Strong customer service Work as a team Relationship building Will mold them on the technology-side internally SaaS Machine learning, artificial intelligence Data analytics and big data Ability to see the “big picture” and entire scope of a project Coding and programming Diversity – gender and ethnicity Nontraditional user interfaces Intelligent interfaces</p>
<p>3. Technology skills a graduate should have</p>	<p>Proprietary systems Agile methodology Project management tools Microsoft Office suite Different platforms CSS, Java, Python</p>

	<p>Angular</p> <p>Web analytics</p> <p>Education vs. training – the complexity of the problem</p>
4. Soft skills a graduate should have	<p>Project management</p> <p>Interpersonal communication</p> <p>Relationship building</p> <p>Problem-solving</p> <p>Working in a team</p> <p>Global teams</p> <p>Communication skills in all formats (web, video, audio, oral, written)</p> <p>Communicate in nontechnical fashion</p> <p>Networking skills</p> <p>Able to accept constructive feedback</p> <p>Student groups</p> <p>Capstone projects</p>
5. Skills and experience that have helped a graduate in finding a position	<p>Relationship building</p> <p>Work with customers in difficult situations</p> <p>Ready to learn</p> <p>Internships</p> <p>Shadowing in a technology company</p> <p>Volunteering/community involvement</p> <p>Opportunities through SoIC's Career Services Office</p> <p>More Q&A in employer sessions</p>
6. Initial PLOs to be retained	<p>Analyze impact of IT</p> <p>Collaborative teamwork</p> <p>Technical knowledge</p> <p>Social dynamics and IT</p> <p>Domain-specific critical thinking and problem-solving skills</p> <p>Collaborative teamwork</p> <p>Professional ethics</p>

Revised Program-Level Learning Outcomes (Revision 2, 10-1-2018)

A. Foundations of Informatics and Computing

1. Explain the fundamentals of computer hardware and software
2. Apply knowledge and skills of logic and discrete mathematics
3. Explain the concepts of statistics and probability
4. Describe data and information representation
5. Select appropriate software to manage information technology projects
6. Illustrate principle and process of interaction design
7. Interpret policies and procedures related to data and information governance
8. Evaluate approaches to information privacy and security

B. Problem Solving and Critical Thinking

1. Use problem-solving techniques to design program algorithms, including pseudocode and flowcharts
2. Explain programming concepts of procedural and object-oriented programming
3. Create computer programs in one or more programming language at the intermediate-to-advanced level of proficiency
4. Appraise the benefits of service-oriented architecture (SOA)
5. Develop insights from data and apply them to address problems and explore opportunities

C. Data Studies and Analytics

1. Apply analytical methods for knowledge and pattern discovery and data analysis
2. Evaluate various data mining and machine learning algorithms
3. Create effective visualizations to analyze and communicate data
4. Communicate insights derived from data

D. Analysis and Design of Information Systems

1. Acquire fundamental concepts of software architecture
2. Develop user requirements
3. Evaluate and create interfaces by applying user experience design theories, terms, principles, and methods
4. Define terms and explain principles essential to design of IT and computing systems
5. Design dynamic, data-driven web applications
6. Design large, complex multilayered information systems with software design patterns
7. Design web service consumers and producers in service-oriented architectures

E. Social Dynamics of Informatics and Information Technology

1. Analyze the social, cultural, and organizational settings in which IT solutions will be deployed to achieve successful implementation
2. Interpret major societal trends affecting the development and deployment of technology, such as access, privacy, intellectual property, security, and equity
3. Analyze the impact of IT on individuals, groups, and organizations at local and global levels
4. Articulate the business considerations of technical knowledge

F. Professional Skills and Ethics

1. Create a personal code of ethics
2. Articulate principles for resolving ethical conflicts
3. Support the ethical and appropriate use of technology
4. Articulate legal and ethical issues when using the intellectual contributions of others
5. Interpret constructive feedback
6. Demonstrate networking skills for personal and professional improvement
7. Communicate IT concepts orally and in writing to nontechnical audiences
8. Work collaboratively as part of a team, including global teams

Domain Specific Knowledge and Skills*

*See specific learning outcomes for individual certificates, areas of specialization, minors, etc.

Evaluation and Dissemination of Results

The project directors presented their work, including their methodology and findings, at the 19th Annual Midwest Scholarship of Teaching and Learning (SoTL) Conference on April 6, 2018, held at the Indiana University–South Bend campus. Their proposal to present at the conference was selected after blind peer review. The presentation was well-received, and the audience also suggested that recent alumni be contacted to provide additional input into the PLOs.

Obstacles and Challenges

Four faculty members participated in the survey. This is primarily because four full-time faculty members are responsible for teaching a substantial percentage of the core courses in the B.S. in Informatics program (and two of these faculty members are the project directors). Moreover, even for some of the elective courses, our faculty members are new to the university or are adjunct, with specialized expertise and did not necessarily have a broad overview of informatics as a discipline.

We had difficulty obtaining interviews with representatives from industry in part because informatics is an interdisciplinary field and does not lend itself to a targeted set of employers or narrow career paths. Representatives from industry indicated that their companies will “on board” new employees with the technical skills they need for their roles, so the feedback from these interviewees tended to focus more on the “soft skills” that are difficult to teach and to assess.

Finally, what we were trying to do in our revisions of the PLOs was to look into the future, rather than reflect the current state of the field of informatics.

Intended Use of Findings for Program Improvement

The project directors will assist faculty in incorporating the revised PLOs into their syllabi and in crafting linkages with assignments that will assess whether courses contribute to helping students achieve these revised PLOs.

The findings will allow SoIC to produce a more comprehensive and data-driven report for the next Program Review and Assessment process for the B.S. in Informatics. The results will help to determine how existing courses should be modified to cover one or more of the PLOs, including assessment mechanisms, and inform the development of new courses.

The revised PLOs are especially important in providing documentation for SoIC’s recruiters, academic advisors and development staff to use in providing a clearer and more compelling picture of the emerging discipline of Informatics and the career opportunities that a B.S. degree in this field offers.

References

1. Program-Level Outcomes (Learning Outcomes), Informatics Bachelor of Science (2016). Retrieved from <http://soic.iupui.edu/undergraduate/degrees/informatics>
2. Principles of Undergraduate Learning (PULs) (1997). Retrieved from <http://due.iupui.edu/Undergraduate-Curricula/General-Education/Principles-of-Undergraduate-Learning>
3. Haviland, D., Turley, S., & Shin, S.H. (2011). Changes over time in faculty attitudes, confidence, and understanding as related to program assessment. *Issues in Teacher Education*, 20(1), 69–84.
4. Hubball, H., & Poole, G. (2008). Program-level assessment of learning outcomes in a faculty certificate program on the Scholarship of Teaching and Learning. *Collected Essays on Learning and Teaching*, 1, 105–109.
5. Akinde, O.A. (2015). A pilot study of students' learning outcomes using didactic and Socratic instructional methods: An assessment based on Bloom's taxonomy. *Educational Research and Reviews*, 10(21), 2821–2833.
6. Attaway, A.N., Chandra, S., Dos Santos, B.L., Thatcher, M.E., & Wright, A.L. (2011). An approach to meeting AACSB assurance of learning standards in an IS core course. *Journal of Information Systems Education*, 22(4), 355–366.
7. Downey, J.P., McMurtrey, M.E., & Zeltmann, S.M. (2008). Mapping the MIS curriculum based on critical skills of new graduates: An empirical examination of IT professionals. *Journal of Information Systems Education*, 19(3), 351–363.
8. Keshavarz, M. (2011). Measuring course learning outcomes. *Journal of Learning Design*, 4(4), 1–9.
9. Lam, B.H., & Tsui, K.T. (2013). Examining the alignment of subject learning through outcomes and course curricula through curriculum mapping. *Australian Journal of Teacher Education*, 38(12), 97–119.
10. Meyers, S.C., Nelson, M.A., & Stratton, R.W. (2011). Assessment of the undergraduate economics major: A national survey. *The Journal of Economic Education*, 42(2), 195–199.
11. Veltri, N.F., Webb, H.W., Matveev, A.G., & Zapatero, E.G. (2011). Curriculum mapping as a tool for continuous improvement of IS curriculum. *Journal of Information Systems Education*, 22(1), 31–42.
12. Cavanaugh, J.K., & Jacquemin, S.J. (2015). A large sample comparison of grade based student learning outcomes in online and face-to-face courses. *Online Learning*, 19(2), 25–32.
13. IUPUI DataWorks (2016). Retrieved from <https://dataworks.iupui.edu/>
14. Original Lilly Conference on College Teaching: Evidence-Based Learning and Teaching (2016). Retrieved from <http://celt.miamioh.edu/lillycon/>
15. Midwest Annual SoTL Conference, University Center for Excellence in Teaching (2016). Retrieved from <https://www.iusb.edu/ucet/events/sotl/>
16. AAEEBL (The Association for Authentic, Experiential and Evidence-Based Learning) (2016). Retrieved from <http://www.aeebl.org/>
17. IUPUI ScholarWorks (2016). Retrieved from <https://scholarworks.iupui.edu/>

Appendix A: Initial Learning Outcomes for the B.S. in Informatics

Graduates of the Informatics undergraduate program will demonstrate expertise in the following core competencies essential to success as an informatics, computing, and information technology professional:

Technical Knowledge

- Demonstrate knowledge and skills in the mathematical and logical foundations of informatics, data representation, models, structures, and informatics-centric management
- Define terms and explain basic principles essential to the design and development of IT and computing systems
- Acquire fundamental concepts and skills in software architectures and the development of information systems

Social Dynamics of Informatics and Information Technology

- Understand and apply major societal trends affecting the development and deployment of modern-day IT, such as access, privacy, intellectual property, security and others
- Critically analyze the impact of IT on individuals, groups and organizations at local and global levels
- Apply a user-centered approach to interaction design and product usability, including techniques for quantitative and qualitative testing of interface and interaction design
- Utilize digital tools to communicate with a range of audiences
- Analyze the social, cultural and organizational settings in which IT solutions will be deployed to achieve successful implementation

Domain-specific* Critical Thinking and Problem-Solving Skills

**Domains are areas of specialization that may include business, science, the arts or humanities.*

- Define terms and explain basic principles, concepts and theories from another domain or discipline in which IT skills will be applied
- Deploy IT resources in the context of another domain and/or discipline
- Synthesize, analyze and conceptualize information and ideas from multiple sources and perspectives
- Evaluate data, arrive at reasoned conclusions and solve challenging problems

Collaborative Teamwork

- Select and effectively utilize oral, written, visual and quantitative communication skills within the context of an interdisciplinary team
- Identify and demonstrate the skills, behaviors and attitudes necessary to function as an effective team member, including working cooperatively with diverse group members

- Acquire the skills to initiate, manage and execute an IT project
- Articulate legal and ethical issues when using the creative work of others; respect the intellectual property of others

Professional Ethics and Development

- Create a personal code of ethics; articulate principles for resolving ethical conflicts
- Commit to a regular program of continuing education and lifelong learning that is independent of employer sponsorship
- Participate in professional organizations that promote responsible computing and service to society

Appendix B: Syllabus Review

Course	Syllabus Year, Semester	POs Included	POs Tied to Assignments
INFO I100	2016, Fall	Y	N
INFO I201	2017, Spring	N	N
INFO I202	2017, Spring	N	N
INFO I210	2016, Fall	N	N
INFO I211	2017, Spring	N	N
INFO I223	2016, Fall	Y	Y
INFO I350	2016, Fall	N	N
INFO I402	2017, Spring	N	N
INFO I453	2017, Spring	N	N
INFO I270	2017, Spring	N	N
INFO I275	2016, Fall	N	N
INFO I303	2016, Fall	N	N
INFO I305	2016, Fall	N	N
INFO I308	2017, Spring	N	N
INFO I330	2017, Spring	Y with link to online document	Y
INFO I421	2016, Fall	N	N
INFO I425	Has not been offered yet		
NEWM N220	2017, Spring	N	N
NEWM N299	2017, Spring	N	N
Total	19	Yes: 3 No: 15 N/A: 1	Yes: 2 No: 16 N/A: 1

Appendix C: Survey Questionnaire of Faculty Members

1. What kinds of skills and knowledge are you looking for in thinking about how to prepare students as part of their B.S. in Informatics degree?
2. What are the short- and long-term trends that you are seeing in IT, especially as they relate to Informatics?
3. What specific technology skills should a graduate of the B.S. in Informatics program have?
4. What “soft skills” should a graduate of the B.S. in Informatics program have?
5. What kinds of skills, knowledge and experiences have helped students be successful in finding a position in the field of Informatics?
6. Of the current PLOs, which should be retained and how can they be assessed or measured?
7. Any other information or insights that you would like to share?

Master of Science in Applied Data Science
Program Review and Assessment Committee Report
Department of Human-Centered Computing
Indiana University School of Informatics and Computing

1. Introduction

This report focuses on the Master of Science in Applied Data Science Program under the IUPUI School of Informatics and Computing. The program currently has three specializations: Applied Data Science, Sports Analytics, and User Experience Design. This program, founded in 2016 as Master of Science in Informatics, is housed in the Department of Human-Centered Computing.

Applied Data Science students learn how to apply data mining and machine learning methods to extract new insights from large datasets. Students also learn how to represent information visually so that it can be communicated effectively to both a technical and lay audience. The program exposes students to industry-standard systems for data management and high-throughput storage and analysis. It includes seven faculty who collectively represent strong research and education expertise in

- Statistics, data mining, and machine learning
- Data visualization and human-computer interaction
- Bioinformatics and computational biology
- Information and library sciences

1.1. Program overview

The M.S. in Applied Data Science is a 30-credit-hour program that integrates topics in informatics, databases, statistics, data mining and visualization, web development, and project management or research design. The program prepares graduates for a career in private industry or academia, or for admission to the Ph.D. in Data Science. The program can be completed in three semesters and a summer, and most students complete the program in 1.5 years.

1.2. Current enrollment

There are 34 students currently in the program, which has picked up a very healthy growth trend of approximately 16 new students for fall 2018, despite the fact that full-time students are no longer guaranteed assistantships, and with a cohort of prospective students that is expected to grow.

2. Program Processes

2.1. Program content

2.1.1. Program structure

The program comprises eight core courses, which are designed to provide students with preparation in the following foundational topics:

- **Statistics and statistical machine learning:** These together constitute the central pillar of data science. Students learn how to apply statistical techniques (e.g., hypothesis testing, linear regression, neural networks) to solve data-driven problems. Additionally, they also gain exposure to the mathematical theory that underpins those tools.

- **Client-server and web application development:** Students learn how to write client–server and “full-stack” applications to query databases, analyze data, and present the resulting information on the web.
- **Human-centered data science:** Students learn how to design human-in-the-loop systems that allow people to see, analyze, and interact with data in real time. This currently includes methods for data visualization and perceptual presentation of data.
- **Large-scale data systems:** students learn to use and apply industry-standard tools for managing the storage and analysis of massive datasets. This includes cloud computing, relational and NoSQL databases, and high-throughput data workflows.
- **Long-term data management and the data lifecycle:** Students learn fundamental techniques for curating sustainable and secure databases and the ethics of data stewardship and preservation.
- **Bioinformatics and biostatistics:** because data scientists often rely on their domain knowledge to solve problems, many of the courses in the program use biological datasets as examples. This allows students to work with real-world datasets through assignments and projects. Additionally, it allows the program to leverage the teaching and research expertise of our faculty, many of whom lead world-class research programs in bioinformatics and genomics.

In addition to the eight courses, the program includes two elective courses. The elective courses enable students to gain depth and breadth expertise in areas of professional or academic interest. This currently includes

- Information policy and data ethics
- Data archiving and long-term storage
- Image analysis and computer vision
- Advanced statistical topics
- Research
- Professional internships (see below)

2.1.2. How has the program curriculum responded to new directions in the discipline?

The MS in Applied Data Science is a new program, established in 2016. Yet, given the dynamic and continuously evolving landscape of the data science field, the program’s plan of study has undergone a revision (for fall 2019 incoming students), with two new components added:

- Deep Learning Neural Networks (INFO-H 518): This course addresses an emerging class of machine learning techniques, deep neural networks. These techniques are capable of solving difficult problems and have applications in a large number of domains, including face and object recognition, medical diagnosis (e.g., of cancerous histologies). The course was first offered as an elective in spring 2018 and, given the high demand for deep learning skills in industry, will be offered every spring.
- Students now have the option of enrolling in up to the credit hours professional internships as an elective. The student is required to work a minimum of 45 hours per credit hour and submit weekly work journals to advisors (see section 2.2.2).

2.2. Student support and mentoring

2.2.1. How and when are research advisors selected for graduate students?

For course advising, the advising to the students operates at two levels:

1. *Graduate program advising and orientation*: The Graduate Program Coordinator engages with students to orient them from admission to graduation to fulfill the necessary verifications and requirements to maintain academic standing, including grade requirements, full-time/part-time status, support for international students, and liaison with the Office of International Affairs (OIA), degree and course transfers, credit transfers, and leave of absences.
2. *Plan of study advising*: The Program Director provides general guidance to the students on the plan of studies, organization of the course load for each semester, selection of the electives, and suggestions to contact faculty with common interests or related research or industrial projects.

For research advising, the following process is followed:

- Upon admission, faculty review a self-reported technical and research skillset and the personal statement in the students' application material.
- Based on this information, by the beginning of the first semester, faculty express to the assistant to the Chair their preference for engaging students in research projects based on their research agenda and the student's interests and skillset.
- Faculty and students are matched, with the opportunity to reassess performance and fit each semester.

2.2.2. Internships and professional development

Starting in fall 2018, students will be encouraged to do internships that are relevant to their professional career. Students are able to take advantage of experiential learning opportunities as elective credits throughout the program for up to three credit hours. It is on these internships they are able to integrate knowledge and theory learned in the classroom with practical application and development of skills in the professional under the supervision of a mentoring supervisor and course instructor.

The interning student is required to turn in weekly journal entries and a written report describing their activities and accomplishments.

2.2.3. Processes to help graduate students learn to teach and selection of teaching assistants

Admitted master students are vetted and selected to serve as Research or Teaching Assistants based on the interests of the students, their self-report skillset and the needs of the department. Since the first semester of admission, full-time MS students may apply for a teaching or research assistantship (funded by the department) to assist faculty in undergraduate Informatics courses and a faculty mentor (the course instructor) who directly supervises their work. Most Applied Data Science MS students are assigned a Graduate Assistantship of 5, 10, or 20 hours per week. For Teaching Assistants, the hours are mainly determined by the number of undergraduate students in the courses and the skillset of the teaching assistant. The performance of teaching assistants is monitored by faculty in assigning them increasingly challenging teaching roles (from grading, to class supervision, preparation and logistics, to student interaction, tutoring, coaching, and lecturing) and providing feedback to the student and to the Chair. Students are encouraged to take advantage of the Center for Teaching and Learning on campus to hone their communication and teaching skills.

3. Learning Outcomes

Program-level learning outcomes were developed based on competing programs, such as U.C., Berkeley's Master of Information and Data Science and Georgia Tech's Master of Science in Data Science, and the vision that the program should span the data science pipeline, which includes data acquisition, curation,

distributed storage, management, analysis, and visualization. Courses for the program were selected and additional courses were created to meet these learning outcomes. Program-level and course-level learning outcomes were brought into exact alignment through negotiation between the Program Director and individual instructors.

The program-level learning outcomes are broadly divided into six main competencies:

1. Statistical learning, data analytics, and deep learning models.
2. Data management, infrastructure, and the data science lifecycle.
3. Client–server application development.
4. Management of large-scale data stores and cloud-based computing systems.
5. Interactive data visualization and human-centered data science.
6. Ethical and professional management of informatics projects.

The specific learning outcomes under each competency area are outlined below.

3.1. Competency in statistics, data analytics, and deep learning models

	Primary Course Outcome Number	Revised Bloom's Taxonomy
1. Differentiate between research fields, theoretical concepts, epistemologies, and qualitative and quantitative methods.	INFO I501-LO1	4
2. Analyze critically and speak publicly about field-specific scholarly research, projects executed in class, and data management issues.	INFO I501-LO2	4
3. Design, implement, test, and debug extensible and modular programs involving control structures, variables, expressions, assignments, I/O, functions, parameter passing, data structures, regular expressions, and file handling.	INFO I501-LO3	5, 6
4. Analyze computational complexity in algorithm development.	INFO I501-LO4	4
5. Investigate research questions and designs by loading, extracting, transforming, and analyzing data from various sources.	INFO I501-LO5	5, 6
6. Test hypotheses and evaluate reliability and validity.	INFO I501-LO6	4
7. Implement histograms, classifiers, decision trees, sampling, linear regression, and projectiles in a scripting language.	INFO I501-LO7	3
8. Decompose and simulate systems to process data using randomness.	INFO I501-LO8	6
9. Employ supervised and unsupervised machine learning for functional approximation and categorization.	INFO I501-LO9	4
10. Analyze datasets with supervised learning methods for functional approximation, classification, and forecasting and unsupervised learning methods for dimensionality reduction and clustering.	INFO H515-LO1	4
11. Explore, transform, and visualize large, complex datasets with graphs in R.	INFO H515-LO2	5, 6
12. Solve real-world problems by adapting and applying statistical learning methods to large, complex datasets.	INFO H515-LO3	4, 5
13. Identify, assess, and select appropriately among statistical learning methods and models for solving a particular real-world problem, weighing their advantages and disadvantages.	INFO H515-LO4	4, 5
14. Write programs to perform data analytics on large, complex datasets.	INFO H515-LO5	6

15. Analyze datasets from case studies in informatics fields (e.g., digital media, HCI, health informatics, bioinformatics, and business intelligence).	INFO H515-LO6	4
16. Solve problems in linear algebra, probability, optimization, and machine learning	INFO H518-LO1	4
17. Evaluate, in the context of a case study, the advantages and disadvantages of deep learning neural network architectures	INFO H518-LO2	5
18. Implement deep learning models in Python using the PyTorch library and train them with real-world datasets.	INFO H518-LO3	3, 6
19. Design convolution networks for handwriting and object classification from images or video.	INFO H518-LO4	5, 6
20. Design recurrent neural networks with attention mechanism for natural language classification, generation, and translation	INFO H518-LO5	5, 6
21. Evaluate the performance of different deep learning models (e.g., with respect to the bias-variance trade-off, overfitting, and underfitting, estimation of test error).	INFO H518-LO6	5
22. Perform regularization, training, optimization and hyperparameter selection on deep models.	INFO H518-LO7	4
23. Analyze a deep learning model's hardware node and GPU scalability in preparation for deployment.	INFO H518-LO8	4

3.2. Competency in data management, infrastructure, and the data science lifecycle

	Primary Course Outcome Number	Revised Bloom's Taxonomy
1. Design and implement relational databases using tables, keys, relationships, and SQL commands to meet user and operational needs.	LIS S511-LO1	3, 6
2. Diagram a relational database design with entity-relationship diagrams (ERDs) using crow's foot notation to enforce referential integrity.	LIS S511-LO2	4
3. Evaluate tables for compliance to third normal form and perform normalization procedures on noncompliant tables.	LIS S511-LO3	5
4. Write triggers to handle events and create views to enforce business rules within a relational database.	LIS S511-LO4	4, 6
5. Demonstrate an understanding of the data lifecycle, including data curation, stewardship, preservation, and security.	LIS S511-LO5	3
6. Evaluate the social and ethical implications of data management.	LIS S511-LO6	5

3.3. Competency in client–server application development

	Primary Course Outcome Number	Revised Bloom’s Taxonomy
1. Design and implement client–server applications that solve real-world problems.	NEWM N510-LO1	6
2. Create well-formed static and dynamic webpages using current versions of PHP, HTML, CSS, and JavaScript or their equivalents.	NEWM N510-LO2	6
3. Implement the model-view-controller software pattern in web and mobile user interfaces.	NEWM N510-LO3	3, 6
4. Apply client-side and server-side programming skills including design, coding, implementation, and integration with relational databases.	NEWM N510-LO4	3–6
5. Extract data from JavaScript Object Notation (JSON) and Extensible Markup Language (XML) documents.	NEWM N510-LO5	3
6. Transmit objects between the browser and server by converting them into JSON.	NEWM N510-LO6	3
7. Evaluate a given web application based on different criteria such as structure, dynamics, security, embedded systems, and interactivity.	NEWM N510-LO7	4, 5
8. Diagram the phases of the secure software development lifecycle.	NEWM N510-LO8	3, 6
9. Demonstrate the techniques of defensive programming and secure coding.	NEWM N510-LO9	4, 6
10. Design user-friendly web and mobile interfaces.	NEWM N510-LO10	6

3.4. Competency in the management of large-scale data stores and cloud computing

	Primary Course Outcome Number	Revised Bloom’s Taxonomy
1. Research the main concepts, models, technologies, and services of cloud computing, the reasons for the shift to this model, and its advantages and disadvantages.	INFO H516-LO1	5
2. Examine the technical capabilities and commercial benefits of hardware virtualization.	INFO H516-LO2	4
3. Analyze tradeoffs for data centers in performance, efficiency, cost, scalability, and flexibility.	INFO H516-LO3	4
4. Evaluate the core challenges of cloud computing deployments, including public, private, and community clouds, with respect to privacy, security, and interoperability.	INFO H516-LO4	5
5. Create cloud computing infrastructure models.	INFO H516-LO5	6
6. Demonstrate and compare the use of cloud storage vendor offerings.	INFO H516-LO6	4
7. Develop, install, and configure cloud-computing applications under software-as-a-service principles, employing cloud-computing frameworks and libraries.	INFO H516-LO7	3, 4
8. Apply the MapReduce programming model to data analytics in informatics-related domains.	INFO H516-LO8	3
9. Enhance MapReduce performance by redesigning the system architecture (e.g., provisioning and cluster configurations).	INFO H516-LO9	4–6
10. Overcome difficulties in managing very large datasets, both structured and unstructured, using nonrelational data storage and retrieval (NoSQL), parallel algorithms, and cloud computing.	INFO H516-L10	4–6

3.5. Competency in interactive data visualization and human-centered data science

	Primary Course Outcome Number	Revised Bloom's Taxonomy
1. Assess the purpose, benefits, and limitations of visualization as a human-centered data analysis methodology.	INFO H517-LO1	5
2. Conceptualize and design effective visualizations for a variety of data types and analytical tasks.	INFO H517-LO2	6
3. Implement interactive visualizations using modern web-based frameworks.	INFO H517-LO3	3, 4, 6
4. Evaluate critically visualizations using perceptual principles and established design guidelines.	INFO H517-LO4	5
5. Conduct independent research on a range of theoretical and applied topics in visualization and visual analytics.	INFO H517-LO5	5-6

3.6. Competency in the ethical and professional management of informatics projects

	Primary course-Outcome number	Revised Bloom's Taxonomy
1. Apply project management methods to overcome the complexities of informatics projects.	INFO B505-LO1	3
2. Plan informatics projects, setting their scope and assigning team members appropriately to roles.	INFO B505-LO2	6
3. Apply to informatics projects time management concepts, such as network diagrams, CPM, and PERT.	INFO B505-LO3	3
4. Apply cost management and budgeting principles.	INFO B505-LO4	3
5. Manage unanticipated changes in informatics projects.	INFO B505-LO5	4-6
6. Perform risk analysis by means of quantitative and qualitative methods.	INFO B505-LO6	4, 5
7. Employ both "hard" and "soft" skills in leading a project team.	INFO B505-LO7	3
8. Use project management software effectively.	INFO B505-LO8	3
9. Apply communication, negotiation, and group decision-making abilities in team projects.	INFO B505-LO9	3
10. Demonstrate ethical and professional behavior in response to ethically challenging situations.	INFO B505-L10	2-6

4. Evidence of meeting learning outcomes

4.1. Direct measures

4.1.1. Methodology

To evaluate how well students are meeting the program's intended learning outcomes (LOs), we conducted quantitative analysis of student grades for all assignments from seven core courses in the program. Anonymized student grades were collected from instructors for courses taught in fall 2017 and spring 2018. The grades were broken down by individual assignment, project, exam, or quiz. For example, if a course contained three assignments with 30 students enrolled, we collected the grade received by every enrolled student for each of the three assignments individually (i.e., a total of $30 \times 3 = 90$ assessment points). Grades were renormalized so that they range from 0 to 100 points to allow meaningful comparison between courses. We then asked course instructors to map each assignment to program-level LOs. A total of 1,382 assessments were collected using this method. The number of course enrollments represented in this sample is 199. Note that this number does not represent unique students, because students typically enroll in multiple courses simultaneously. The collected dataset covers 48 of

the program's 64 learning outcomes (75%). Data could not be collected on 16 LOs due to lack of response from corresponding instructors at the time of writing.

4.1.2. Results

We computed grade distributions from the above data. These distributions represent instructor assessment of students' mastery of each program-level LO. The distributions are plotted as histograms in the following table, broken down by competency area then by course. The red line within the histogram plots represent a 80-point grade, the program's threshold for passing. We also include the percentage of students (shown to the right of the histogram) who are considered to have mastered the intended outcome, having achieved a passing grade (≥ 80 points). The percentile is color coded in green ($\geq 70\%$), orange ($\geq 50\%$), or red ($< 50\%$).

Competency in statistics, data analytics, and deep learning models

		grade distribution	% student meeting learning outcome
INFO 1501	LO1: Differentiate between research fields, theoretical concepts, epistemological, and qualitative and quantitative methods.		96.9%
	LO2: Analyze critically and speak publicly about field-specific scholarly research, projects executed in class and data management issues.		100%
	LO3: Design, implement, test, and debug extensible and modular programs involving control structures, variables, expressions, assignments, I/O functions, parameter passing, data structures, regular expressions, and file handling.		93.8%
	LO4: Analyze computational complexity in algorithm development.		76.6%
	LO5: Investigate research questions and designs by loading, extracting, transforming, and analyzing data from various sources.		89.1%
	LO6: Test hypotheses and evaluate reliability and validity .		68.8%
	LO7: Implement histograms classifiers, decision trees, sampling, linear regression, and projectiles in a scripting language.		98.4%
	LO8: Decompose and simulate systems to process data using randomness.		51.6%
	LO9: Employ supervised and unsupervised machine learning for functional approximation and categorization.		71.9%

INFO H518

	grade distribution	% student meeting learning outcome
LO1: Solve problems in linear algebra, probability , optimization, and machine learning.		100%
LO2: Evaluate, in the context of a case study , the advantages and disadvantages of deep learning neural network architectures and other approaches.		18.2%
LO3: Implement deep learning models in Python using the PyTorch library and train them with real-world datasets.		27.3%
LO4: Design convolution networks for handwriting and object classification from images or video.		81.8%
LO5: Design recurrent neural networks with attention mechanisms for natural language classification, generation, and translation.		81.8%
LO6: Evaluate the performance of different deep learning models (e.g., with respect to the bias-variance trade-off, overfitting and underfitting, estimation of test error).		100%
LO7: Perform regularization, training optimization, and hyperparameter selection on deep models.		18.2%
LO8: Analyze a deep learning model's hardware node and GPU scalability in preparation for deployment.		100%

INFO H515

	grade distribution	% student meeting learning outcome
LO1: Analyze datasets with supervised learning methods for functional approximation, classification, and forecasting and unsupervised learning methods for dimensionality reduction and clustering.		71.0%
LO2: Explore, transform, and visualize large, complex datasets with graphs in R.		93.5%
LO3: Solve real-world problems by adapting and applying statistical learning methods to large, complex datasets.		96.8%
LO4: Identify, assess, and select among statistical learning methods and models for solving a particular real-world problem, weighing their advantages and disadvantages.		32.3%
LO5: Write programs to perform data analytics on large, complex datasets in R.		51.6%
LO6: Analyze datasets from case studies in informatics fields (e.g., digital media, HCI, health informatics, bioinformatics, and business intelligence).		90.3%

Competency in data management, infrastructure, and the data science lifecycle

		grade distribution	% student meeting learning outcome
LIS S511	LO1: Design and implement relational databases using tables, keys, relationships, and SQL commands to meet user and operational needs.		95.2%
	LO2: Diagram a relational database design with entity-relationship diagrams (ERDs) using crow's foot notation to enforce referential integrity .		95.2%
	LO3: Evaluate tables for compliance to third normal form and perform normalization procedures on noncompliant tables.		95.2%
	LO4: Write triggers to handle events and enforce business rules and create views within a relational database.		100%
	LO5: Demonstrate an understanding of the data lifecycle, including data curation, stewardship, preservation, and security .		100%
	LO6: Evaluate the social and ethical implications of data management.		71.4%

Competency in client-server application development

		grade distribution	% student meeting learning outcome
NEWM N510	LO1: Design and Implement client-server applications to solve real-world problems.		100%
	LO2: Create well-formed static and dynamic webpages using current versions of PHP , HTML, CSS, and JavaScript, or their equivalents.		54.5%
	LO3: Implement the model-view controller software pattern in web and mobile user interface.		81.8%
	LO4: Apply client-side and server-side programming skills, including design, coding, implementation, and integration, with relational databases.		81.8%
	LO5: Extract data from JavaScript Object Notation (JSON) and Extensible Markup Language (XML) documents.		100%
	LO7: Evaluate a given web application based on dif ferent criteria, such as infrastructure, dynamics, security , embedded systems, and interactivity .		90.9%
	LO8: Diagram the phases of the secure software development lifecycle.		81.8%
	LO9: Demonstrate the techniques of defensive programming and secure coding.		100%
	LO10: Design user-friendly web and mobile interfaces.		100%

Competency in the management of large-scale data stores and cloud computing

		grade distribution	% student meeting learning outcome
INFO H516	LO1: Research the main concepts, models, technologies, and services of cloud computing, the reasons for the shift to this model, and its advantages and disadvantages.		6.7%
	LO7: Develop, install, and configure cloud-computing applications under software-as-a-service principles, employing cloud-computing frameworks and libraries.		56.7%
	LO8: Apply the MapReduce programming model to data analytics in informatics-related domains.		76.7%
	LO9: Enhance MapReduce performance by redesigning the system architecture (e.g., provisioning and cluster configurations).		90.0%
	LO10: Overcome difficulties in managing very large datasets, both structured and unstructured, using non-relational data storage and retrieval (NoSQL), parallel algorithms, and cloud computing.		50.0%

Competency in interactive data visualization and human-centered data science

		grade distribution	% student meeting learning outcome
INFO H517	LO1: Assess the purpose, benefits, and limitations of visualization as a human-centered data analysis methodology .		96.8%
	LO2: Conceptualize and design effective visualizations for a variety of data types and analytical tasks.		90.3%
	LO3: Implement interactive visualizations using modern web-based frameworks.		100%
	LO4: Critically evaluate visualizations using perceptual principles and established design guidelines.		100%
	LO5: Conduct independent research on a range of theoretical and applied topics in visualization and visual analytics.		100%

4.1.3. Discussion

Analysis of student grades reveal that 37 LOs (or 77.1% of the evaluated LOs) show evidence of good student achievement, with over 70% of student obtaining a passing grade in exams, assignments, and/or projects. Additionally, 6 (12.5%) LOs demonstrate moderate success (with over 50% of students achieving a passing grade). Lastly, 5 (10.4%) LOs have low achievement levels (less than 50% of enrolled students obtaining a passing grade).

Five potentially problematic LOs that were identified are listed (and highlighted in **red** in the assessment table). Course assessments indicate that less than 50% of enrolled students had mastered these LOs:

1. **INFO H516-LO1**: Research the main concepts, models, technologies, and services of cloud computing, the reasons for the shift to this model, and its advantages and disadvantages.
2. **INFO H518-LO2**: Evaluate, in the context of a case study, the advantages and disadvantages of deep learning neural network architectures and other approaches.

3. **INFO H518-LO3:** Implement deep learning models in Python using the PyTorch library and train them with real-world datasets.
4. **INFO H518-LO7:** Perform regularization, training optimization, and hyperparameter selection on deep models.
5. **INFO H515-LO4:** Identify, assess, and select among statistical learning methods and models for solving a particular real-world problem, weighing their advantages and disadvantages.

Six moderately successful outcomes (50%-70% achievement level) indicate potential room for improvement (highlighted in **orange**):

1. **INFO I501-LO6:** Test hypotheses and evaluate reliability and validity.
2. **INFO I501-LO8:** Decompose and simulate systems to process data using randomness.
3. **INFO H516-LO7:** Develop, install, and configure cloud-computing applications under software-as-a-service principles, employing cloud-computing frameworks and libraries.
4. **INFO H516-LO10:** Overcome difficulties in managing very large datasets, both structured and unstructured, using non-relational data storage and retrieval (NoSQL), parallel algorithms, and cloud computing.
5. **NEWM N510-LO2:** Create well-formed static and dynamic webpages using current versions of PHP, HTML, CSS, and JavaScript, or their equivalents.
6. **INFO H515-LO5:** Write programs to perform data analytics on large, complex datasets in R.

4.1.4. Limitations:

Although revealing, the quantitative analysis above has two limitations. First, it relies entirely on student grades. While grades are a good proxy of instructor's assessment of whether students are meeting learning outcomes, grading difficulty and rigor can vary across instructors and classes. Secondly, some assignments span multiple learning goals, which can make it difficult to map student performance to individual learning outcomes for some courses.

We supplement the above quantitative analysis with indirect measures (see section 4.2). In addition, program instructors analyzed student feedback and evaluations, and incorporated those in course revisions (see section 5).

4.2. Indirect evidence of meeting learning outcomes

We conducted a post-graduation job placement survey to collect data about the employment status and annual salaries of graduating students. Since its inception in 2016, the program has graduated 18 MS students. As of June 2018, our survey indicates that 69% (11 students) are already employed in part-time, full-time, or in internship positions. Furthermore, 63% of graduates are full-time employed in within-major positions. Students who were able to secure employment found positions at nationally recognized and in-state employers, including Amazon, Eli Lilly, Cummins Inc., and the Indiana State Department of Health. The median reported salary for graduates was \$72,500 annually.

This preliminary data indicate that majority of graduating students are gainfully employed upon completing the program. Most graduates found positioned that are within-major, consistent with the program's learning outcomes and career aspirations. Additionally, job placements in several Fortune 500 companies (Amazon, Eli Lilly, and Cummins Inc.) suggest that employers recognize the knowledge and skills possessed by our graduating students.

That said, given that our program is still young, the sample of available graduates is somewhat limited. We will therefore continue to monitor employment rates and salary and follow up with our alumni to acquire additional data about their job placement and career progression and satisfaction.

5. Course-level revisions and improvements

The quantitative analysis above was shared with all program instructors. Instructors reflected on a combination of student evaluation and in-class observations by peers or teaching consultants. The following revisions were made to program courses in response.

5.1. Improvement to meet learning outcomes with low student performance

As discussed in section 4.1, five LOs exhibited relatively low student performance, with less than 50% of students obtaining a passing grade. Of those outcomes, three are in INFO-H 518“Deep Learning Neural Networks, one is in INFO-H 516 Applied Cloud Computing for Data Intensive Sciences, and one is in INFO-H 515 Data Analytics. Instructors of the above three courses introduced the following revisions in response.

5.1.1. INFO-H 515 Data Analytics

The course was modified substantially after Fall 2017. In addition to student course evaluations, Prof. Lourens solicited feedback on the course after student presentations to identify areas for course improvement. Students felt the course was too focused on the R statistical programming environment, which did not leave time to cover statistical learning thoroughly enough (INFO H515-LO5). In response, multiple R lectures were removed from the course. The reduction of R materials will allow time for students to focus on and absorb the material on statistical learning. Additionally, it will lead to higher student satisfaction and a more cohesive course structure.

5.1.2. INFO-H 516 Applied Cloud Computing for Data Intensive Sciences

This course is currently in its first offering and will benefit from revisions based on student evaluation:

- The most common complaint in student feedback was lack of course materials for some of the course topics. As a result, the students had to rely on lecture slides. Prof. Chakraborty plans to address this by sharing more online resources for each topic and by including more details in the slides. He also plans to record lectures, so that students can revisit the lecture at any time.
- The course incorrectly assumed a strong computer science background. This will be addressed by making the materials more approachable to students from different backgrounds.
- Assessments relied heavily on programming assignments. Quizzes will be added to assess acquisition of theoretical concepts. This is expected to improve achievement for INFO H516-LO1. Furthermore, the grading rubric will be made more transparent.
- In addition, (1) management of the Aspen cluster will be improved so that it does not fail frequently; and (2) quieter students will be engaged by making the class more interactive and having more in-class activities.

5.1.3. INFO-H 518 Deep Learning Neural Networks

This course is in its first offering and will benefit from significant revisions following student evaluation and quantitative analysis of learning outcomes. The following changes will be introduced by Prof. Chakraborty in the course’s spring 2019 offering:

- The assignments will be re-designed so that they are less rigid, giving students more freedom to explore. The new assignments will be more open-ended, which should improve student's ability to generalize and apply concepts learned in the class (INFO H518-LO2).
- The instructor will provide individualized, technical feedback after each assignment and quiz to address student evaluations (while addressing INFO H518-LO3 and INFO H518-LO7).
- The instructor will recruit a more knowledgeable TA.

5.2. Additional course revisions in response to student evaluation and feedback

Program instructors have made (or will make) the following revisions in response to qualitative student evaluations and peer observation of teaching:

5.2.1. NEWM-N 510 Web Database Development

The class has a lecture portion and a lab portion. In the lab portion, students work individually or in groups to complete mini-projects that give them practice implementing the concepts discussed in class. Based on reviews from a consultant (from IUPUI's Center for Teaching and Learning) and the students, this year Prof. Martin-Hammond included more in-class discussion and examples to make the lecture portion more interactive. The instructor also opened each lecture by focusing on specific challenges the students faced with the information and examples presented in reading materials to emphasize the expectation that they are to actively look at materials before class so that they can participate. Also, last year, students worked on projects during the lab portion of the class and were encouraged to work in groups, but many worked individually. On the recommendation of the observer and students, the instructor made the class labs group-based to encourage learning among peers.

5.2.2. LIS-S 511 Database Design

The feedback from the student course evaluations was that the course was not challenging enough with regards to Database Concepts and SQL. Prof. Murillo plans to revise the course to remedy this. The outcome of the revision will be students who are more accurately assessed and better prepared for more advanced courses in the curriculum and employment in industry. Specifically, the following changes will be made:

- Regarding Database Concepts, the instructor will increase the difficulty of the weekly homework by focusing on real-world application-based exercises instead of basic concepts. Conceptual instruction and assessment will be incorporated directly into class activities to ensure that students understand concepts and terminology. Additionally, more hands-on activities will be incorporated into class to ensure that students are ready to accomplish advanced homework exercises.
- Regarding SQL, the instructor plans to increase the difficulty level of this unit are two-fold. First, she will introduce more advanced SQL querying to the course content and homework. Second, she will provide more content regarding MySQL instead of Microsoft Access as the relational database management system. This shift in focus to open-source Unix-based technologies will leave students better prepared for intermediate courses like NEWM-N 510 Web Database Development and advanced courses like INFO-H 516 Applied Cloud Computing for Data Intensive Sciences.

5.2.3. INFO-H 517 Visualization Design, Analysis, and Evaluation

Student evaluation indicate the course workload was heavy. In particular, requiring students to complete 3 projects, 1 homework, and 1 paper presentation was seen as excessive by most students. Additionally, a

number of students felt that the course covered the theoretical materials enough, but less attention was devoted to technical aspects of visualization design (specifically, using the D3 language to create visualizations). In response, Prof. Reda introduced the following changes:

- The number of projects was reduced from 3 to 2: 1 individual project and 1 group project. This should significantly reduce the course workload.
- The instructor added 3 lab-style sessions (to be delivered during regular class time). The sessions will be group-based (to allow peer learning) and will focus on honing students' technical skills and fluency with D3 and JavaScript (the primary tools used to create visualizations). These skills are directly tied to INFO-H 517-LO2 and LO3.

5.2.4. INFO-I 501 Introduction to Informatics

From spring 2015 until an assessment in fall 2017, INFO-I 501 Introduction to Informatics was divided into three 5-week modules: database fundamentals, probability and statistics using Python, and data mining using Weka. The class size gradually increased from 8 in spring 2015 to 42 in fall 2016. Majors include Health Informatics, Data Science, HCI, Computer Science, Public Health, Statistics, Bioinformatics, Nursing, and Business. The course was assessed based on student evaluations, course projects, student majors, and advancements in informatics. Students appreciated applying statistics to data in projects, learning theory and methods, and the midterm debate and autoethnographic essay on a topic using data, usually with clinical or societal relevance.

In response to student requests and changes in the field, the course is now focused on learning Python to retrieve, process, and visualize data. It is often a student's first programming course but simultaneously focuses on applying Python to processing data. Because real-world projects often still use relational databases, we spend the first 2 weeks on learning SQL select and insert. However, we do not teach database design, because various database courses are already available. Based on changes in the field, we added machine learning using *scikit* to apply supervised and unsupervised algorithms to student project datasets. In addition to statistical methods of linear and logistic regressions, we implement clustering and classification. The project presentation now includes peer grading, because student contribution varies. This has improved student participation. Many students have discussed that the course is very intensive with weekly quizzes and assignments, a midterm essay, and project. Thus, we removed an introduction to the subfields of informatics.