# MARYLAND HIGHER EDUCATION COMMMISSION
# ACADEMIC PROGRAM PROPOSAL

## PROPOSAL FOR:

✓ NEW INSTRUCTIONAL PROGRAM

____ SUBSTANTIAL EXPANSION/MAJOR MODIFICATION

____ COOPERATIVE DEGREE PROGRAM

✓ WITHIN EXISTING RESOURCES or ____ REQUIRING NEW RESOURCES

*(For each proposed program, attach a separate cover page. For example, two cover pages would accompany a proposal for a degree program and a certificate program.)*

**Johns Hopkins University**
Institution Submitting Proposal

**Fall 2016**
Projected Implementation Date

| **Master of Science** | **Data Science** |
|---|---|
| Award to be Offered | Title of Proposed Program |

| | **27.0501** |
|---|---|
| Suggested HEGIS Code | Suggested CIP Code |

| **Whiting School of Engineering** | **T.E. Schlesinger, Dean** |
|---|---|
| Department of Proposed Program | Name of Department Head |

| **Philip Tang** | **alo@jhu.edu** | **(410) 516-6087** |
|---|---|---|
| Contact Name | Contact E-Mail Address | Contact Phone Number |

Robert C. Pond **03/15/2016**
Signature and Date — President/Chief Executive Approval

**N/A**
Date — Date Endorsed/Approved by Governing Board

JOHNS HOPKINS
UNIVERSITY

MAR 2 1 2016

March 18, 2016

Dr. James D. Fielder, Jr.
Secretary
Maryland Higher Education Commission
6 N. Liberty Street, 10th Floor
Baltimore, MD 21201

Dear Dr. Fielder:

On behalf of Provost Robert Lieberman, Dean T.E. Schlesinger, and our Whiting School of Engineering, I write to request your review and endorsement of the enclosed proposal for a new **Master of Science in Data Science.**

The Master of Science in Data Science is designed to provide students pursuing careers in data science with a course of study that balances theory with practice, giving them the knowledge and skills necessary to be effective data scientists. The proposed program will be offered onsite and online by the Whiting School's Engineering for Professionals division and is intended for part-time students.

The proposed program is consistent with the mission of the university and with the State of Maryland's goals for postsecondary education. The proposal has been approved by the Whiting School Graduate Committee and the Homewood Academic Council and is fully endorsed by Johns Hopkins University.

A business check (#11478831) for the review of this proposal has been sent to the Commission. Should you have any questions or need further information, please do not hesitate to contact me at (410) 516-6087 or ptang@jhu.edu. Thank you for your continuing support of Johns Hopkins.

Sincerely,

Philip Tang
Vice Provost for Academic Affairs

cc: Dr. Robert C. Lieberman
Dr. T.E. Schlesinger
Dr. Dexter Smith
Dr. Daniel Horn
Dr. Ratna Sarkar
Mr. James Brailer
Ms. Mary Ellen Flaherty
Ms. Jennifer Martin
Mr. Tom McDermott

Enclosures

**The Johns Hopkins University**
**Whiting School of Engineering**
**Proposal for a New Academic Program**

**Master of Science in Data Science**

---

## A. Centrality to institutional mission statement and planning priorities

### 1. Program description and alignment with mission

The Johns Hopkins University Whiting School of Engineering is pleased to submit a proposal for a new Master of Science in Data Science. This new program builds on the existing Master of Science in Applied and Computational Mathematics and the existing Master of Science in Computer Science programs. The JHU Whiting School of Engineering has offered both of these degrees since 1984. Students will have the option to take classes in this program either onsite or online.

The Master of Science in Data Science program will provide students pursuing careers in data science with a course of study that balances theory with practice, giving them the knowledge and skills necessary to enhance their effectiveness in a complex and rapidly evolving technological environment. The emerging discipline of data science is concerned with the theory, analysis, design, and implementation of processes that describe and transform information to discover relationships and insights into complex and ever growing data sets. With roots in applied and computational mathematics and computer science, data science uses formal techniques and methodologies of abstraction to create models that can be automated to solve real-world problems. This unique degree program provides the rigorous technical background and experience to rapidly accelerate a program graduate into the emerging roles in a wide diversity of public and private organizations. Most students will complete the Master of Science in Data Science degree in three to five years. This degree program will be available to students beginning in June 2016.

The mission of The Johns Hopkins University is to educate its students and cultivate their capacity for life-long learning, to foster independent and original research, and to bring the benefits of discovery to the world. In addition, the mission of the JHU Whiting School of Engineering is to provide educational programs of the highest quality that will attract the most qualified and driven students and faculty and will be a world-recognized leader in engineering education, to lead in the creation and dissemination of knowledge, and to translate those educational and research activities into solutions to important societal problems. The proposed degree program aligns with both of these missions as discussed below.

### 2. Alignment with institutional strategic goals

One of the four strategic priorities of the JHU Whiting School of Engineering is to "Educate future leaders by providing students with an innovative and distinctive

education of the highest quality, both at the undergraduate and graduate level, in a diverse and inclusive environment." One of the near-term goals within that priority is to develop a comprehensive suite of contemporary master's degree offerings, for full-and part-time students, with flexible formats that respond to the needs of industry in both the domestic and international markets.

The Johns Hopkins University Engineering for Professionals (JHU-EP) programs in the fields of engineering and applied science are among the oldest and largest part-time programs of their kind in the United States. Administered by the Whiting School of Engineering, this division seeks to meet the lifelong education needs of working professionals in engineering and applied science. JHU-EP offers state-of-the-art courses combined with the convenience, flexibility, and accessibility that make these educational opportunities feasible for working adults.

The educational objective of the Data Science program is to provide students pursuing careers in data science with a course of study that balances theory with practice giving them the knowledge and skills necessary to be effective data scientists.

Graduates in Data Science will be well prepared for specialized jobs solving real world problems involving all aspects of data science from the data pipeline and storage through the statistical analysis and eliciting the story the data has to tell.

## B. Adequacy of curriculum design and delivery to related learning outcomes

### 1. Program outline and requirements

A full course listing with course titles and descriptions is provided in Appendix A. All courses are three (3) credits.

**Admission Requirements**

General admission requirements for master's degree candidates and others seeking graduate status are as follows: applicants must be in the last semester of undergraduate study or hold a bachelor's degree from a regionally accredited college or university.

In addition, applicants for the Master of Science in Data Science must have taken up through multivariate calculus; discrete mathematics; courses in Java or C++, Python, and R; and a course in data structures. Applicants must have received a grade of B- or better in each of the undergraduate prerequisite data science courses.

Applicants who have not taken the prerequisite undergraduate courses may satisfy admission requirements by completing the specified courses with grades of A or B. The program offers the following undergraduate courses, which may be taken as needed to satisfy the prerequisites.

| | |
|---|---|
| 605.101 | Introduction to Python |
| 605.201 | Introduction to Programming Using Java |
| 605.202 | Data Structures |

| | |
|---|---|
| 605.203 | Discrete Mathematics |
| 625.201 | General Applied Mathematics |
| 625.250 | Applied Mathematics I |
| 625.251 | Applied Mathematics II |

**Degree Requirements**

In order to earn a Master of Science in Data Science, 10 courses (30 credits) approved by an adviser must be completed within five years. At least five courses (15 credits) must be from the Applied and Computational Mathematics program, and at least five courses (15 credits) must be from the Computer Science program as specified below. At least two of the data science courses (6 credits) must be 700-level with at least one from Applied and Computational Mathematics and at least one from Computer Science. No more than one course (3 credits) with a grade of C, and no course with a grade lower than C, may be counted towards the degree.

Students will take four required courses (12 credits) and one elective course (3 credits) from each of the Applied and Computational Mathematics track and the Computer Science track.

The required courses are:

| | |
|---|---|
| 605.421 | Foundations of Algorithms |
| 605.441 | Principles of Database Systems |
| 605.448 | Data Science |
| 605.462 | Data Visualization |
| 625.403 | Statistical Methods and Data Analysis |
| 625.415 | Nonlinear Optimization |
| 625.461 | Statistical Models and Regression |
| 625.464 | Computational Statistics |

605.7xx One elective to be selected from:

- 605.741 Distributed Database Systems: Cloud Computing and Data Warehouses[1]
- 605.746 Machine Learning
- 605.748 Semantic Natural Language Processing
- 605.788 Big Data Processing Using Hadoop[5]

625.7xx One elective to be selected from:

- 625.721 Probability and Stochastic Process I
- 625.722 Probability and Stochastic Process II
- 625.725 Theory of Statistics I
- 625.726 Theory of Statistics II
- 625.734 Queuing Theory with Applications to Computer Science
- 625.740 Data Mining
- 625.741 Game Theory

[1] Requires 605.481 Enterprise and Web Computing or equivalent background.

625.743 Stochastic Optimization and Control
625.744 Modeling, Simulation, and Monte Carlo

Students who have been waived from required courses may take additional electives in Data Science, including a capstone or independent study, or up to two outside electives from other Whiting School programs. Students who take outside electives from other programs must meet the specific course and program requirements listed for each course. In the event that the student has transfer courses accepted, they will be considered outside electives.

**2. Educational objectives and student learning outcomes**

The educational objective of the Data Science program is to provide students pursuing careers in either applied mathematics or computer science with a course of study that balances both fields, giving them the knowledge and skills necessary to effectively and competitively respond to the growing demand for data scientists.

Graduates will be prepared to describe and transform information to discover relationships and insights into complex and ever growing data sets. They will use formal techniques and methodologies of abstraction to create models that can be automated to solve real-world problems.

**3. General education requirements**

Not applicable.

**4. Specialized accreditation/certification requirements**

Not applicable.

**5. Contract with another institution or non-collegiate organization**

Not applicable.

**C. Critical and compelling regional or statewide need as identified in the State Plan**

**1. Demand and need for program**

This program accommodates the demand for students wishing to pursue this rapidly growing field and the need from the working professionals already in the JHU-EP program for data science related courses relevant to their technical work environments.

**2. Alignment with the 2013 Maryland State Plan for Postsecondary Education**

The proposed program is well aligned with Maryland Ready, the 2013–2017 Maryland State Plan for Postsecondary Education. The Master of Science in Data Science is intended to prepare highly trained scientists and engineers to work in organizations where they can contribute to the needs of society. The long-term success of JHU-EP programs

for working professionals attests to the quality and effectiveness of these programs. This is consistent with the Goal 1 of the State Plan, "Quality and Effectiveness," which asserts that Maryland will enhance its array of postsecondary education programs to more effectively fulfill the evolving educational needs of its students, the state, and the nation.

Similarly, the proposed program is consistent with Goal 4, "Innovation," which articulates Maryland's aspiration to be "a national leader in the exploration, development, and implementation of creative and diverse education and training opportunities that will align with state goals, increase student engagement, and improve learning outcomes..." By leveraging technology in innovative ways to make JHU-EP offerings more accessible and interactive, candidates can pursue "anytime, anywhere" learning opportunities. Candidates can undertake course-related activities at a time and a location most convenient to them, allowing students to participate in and to complete their program even if their work schedules do not permit regular class attendance or if they move away from the Maryland region, thus also supporting Goal 2, "Access, Affordability, and Completion."

The proposed program is also consistent with Goal 5, "Economic Growth and Vitality," which is centered on supporting a knowledge-based economy through increased education and training. The proposed program will prepare highly qualified local scientists and engineers to contribute to the economic growth and vitality by providing life-long learning to scientists and engineers so they can maintain the skills they need to succeed in the workforce.

## D. Quantifiable and reliable evidence and documentation of market supply and demand in the region and State

### 1. Market demand

In 2012, *Forbes* listed the Washington DC area as the number one area for computer science related work.[2] In November 2013, the job matching service The Ladders,[3] listed data science as number three on the list of the fastest growing high-pay jobs. Additionally, a December 3, 2013, article on Corporate Growth states the "technology that extracts, manages, analyzes and converts data into actionable intelligence are in high demand and attracting a lot of investment dollars."[4]

The State of Maryland makes statewide employment projections based on data from the U.S. Bureau of Labor Statistics. Since the Data Science academic area is so new there is no unique Classification of Instructional Program (CIP) code assigned to it. The CIP code closest to the instructional content for this proposed Master of Science in Data Science program is 27.0501 (Statistics). The Standard Occupational Codes (SOC) associated with this CIP by the Bureau of Labor Statistics are: 11-9121 Natural Sciences Managers; 15-2011 Actuaries; 15-2041 Statisticians; 19-3022 Survey Researchers; and 25-1022 Mathematical Science Teachers, Postsecondary. The Maryland Department of Labor

---

[2] http://www.forbes.com/sites/susanadams/2012/09/20/the-cities-with-the-most-computer-science-jobs/

[3] http://www.theladders.com/press-releases/theladders-releases-new-job-evolution-data--middle-management-titles-are-phasing-out

[4] Corporate Growth http://acgcapitalblog.com/capital-growth/growth-capital-alive-well-washington-dc-metro-area/

Licensing and Regulation (DLLR) projects that long-term employment opportunities for these SOC codes will increase from 2012 to 2022 at an annual rate of 0.3% for Natural Sciences Managers; 0.7% for Actuaries; 1.8% for Statisticians; 1.7% for Survey Researchers; and 1.1% for Mathematical Science Teachers, Postsecondary.

Based on the projected market demand, we expect this degree program to be successful.

**2. Educational and training needs in the region**

Data Science is too new a field to be tracked by the Bureau of Labor Statistics (BLS). It does predict that job opportunities nationwide for the related field of Statistics professionals will grow by 34% from 2014 to 2024. The need for this type of education is also manifested in the numerous ads found in newspapers, websites and trade publications seeking highly trained professionals in data science. Job opportunities for the graduates of this program include positions in corporations and government organizations.

**3. Prospective graduates**

It is anticipated that students and employers will support the proposed Master of Science in Data Science due to its relevance to the metropolitan Baltimore-Washington region.

With four schools in the University of Maryland system, plus several private universities, offering undergraduate degrees in mathematics and computer science in the area, there are hundreds of computer science professionals who could be potential applicants to this program. This is in addition to employees at the many technical organizations of the region who attended school in other parts of the country.

The following table shows the number of graduates over the last five years from Maryland universities with master's degrees for programs with the CIP code of 27.0501 noted in section D.1. above. This data was found in the Maryland Higher Education Commission's graduation trend data base.

| Univ. | Degree | Field | 2010 | 2011 | 2012 | 2013 | 2014 |
|---|---|---|---|---|---|---|---|
| UMBC | Masters | Statistics | 9 | 8 | 9 | 12 | 2 |
| UMCP | Masters | Mathematical Statistics | 4 | 7 | 4 | 5 | 3 |
| UMCP | Masters | Survey Methodology | 11 | 0 | 16 | 5 | 13 |

## E. Reasonableness of program duplication

### 1. Similar programs

Globally, many universities are beginning to offer graduate degrees in data analytics as they attempt to capitalize on the strong interest in industry and government in the process of analyzing and leveraging the data pipeline. Some programs offer specialized degrees, e.g., Social Analytics or Marketing Analytics. Many degrees are offered as specializations within an existing degree program, e.g., Computer Science. A large number are offered under the auspices of a school of business.

Because Data Science is such a new academic area, there is no unique CIP code associated with it. The closest CIP code is that associated with Statistics, as noted above in sections D.1. and D.3. As noted in section D.3. above, there are three master's degrees offered in the State of Maryland having the CIP code of 27-0501 (Statistics). In addition, three other master's degrees were found that are related to data analytics. Comparisons of those programs with the proposed new program are provided below.

The UMBC Master of Science in Statistics includes a traditional track as well as two application tracks, one in environmental statistics and one in biostatistics. The UMCP Master of Arts in Mathematical Statistics is a traditional statistics program. These two programs are broadly-based statistics programs that do not have the focus on data science that the proposed program has. The UMCP Master of Science in Survey Methodology has two concentration options: statistical science and social science. This degree is very focused on a single application area and does not cover the range of topics that the proposed degree does.

The University of Maryland University College (UMUC) offers a Master of Science in Data Analytics, the Robert H. Smith School of Business at the University of Maryland -- College Park (UMCP) offers an MS in Marketing Analytics, and the UMCP College of Information Studies offers a Masters of Information Management with a specialization in Data Analytics. The UMUC program offers a practicum but does not have the mathematical emphasis of the proposed degree. The Smith School of Business program has some emphasis on statistics but is strictly focused on analytics for business marketing purposes. The College of Information Studies program is simply a specialization within an existing MS in Information Management. It emphasizes data management and other management concerns with a minimal mathematics foundation. We know of no other program in Maryland like the JHU-EP Data Science M.S. program that emphasizes equally both the mathematics and computer science aspects of this up and coming field.

### 2. Program justification

The JHU-EP program in Data Science is unique in its engineering emphasis, its equal rigor in both the computer science and mathematical sides of the curriculum, and it is taught by highly-regarded researchers and practitioners. Students seek out the program for the value proposition of JHU quality and part-time convenience. Some schools offer convenience while others offer faculty expertise; the JHU-EP program offers both.

## F. Relevance to Historically Black Institutions (HBIs)

### 1. Potential impact on implementation or maintenance of high-demand programs at HBIs

Web searches of the HBIs in Maryland did not turn up any engineering master's degree offerings in the area of data science.

### 2. Potential impact on the uniqueness and institutional identities and missions of HBIs

By definition, an appropriate student for the proposed Master of Science in Data Science would apply after attending and completing a baccalaureate degree at any undergraduate institution, including any of Maryland's Historically Black Institutions. The proposed program would not directly affect the implementation, maintenance, uniqueness, identity or mission of these institutions.

## G. Evidence of the Principles of Good Practice

See Appendix B for the evidence that this program complies with the Principles of Good Practice noted above. It is possible for some students to complete more than half of the courses online.

The Higher Education Opportunity Act (HEOA) enacted in 2008 requires that an academic institution that offers distance education opportunities to students 1) has a process established to verify that the student who registers is the same student who participates in and completes the offering and receives academic credit for it, 2) has a process established, to verify that student privacy rights are protected, and 3) has a process established that notifies the student about any additional costs or charges that are associated with verification of student identity. In this graduate program, the following actions have been taken to satisfy these requirements: 1) students may only enter the academic website for the online courses they take by providing their unique student ID and password they receive when they are admitted to the program, 2) all FERPA privacy rights are preserved by limiting access very specifically in the University student information system to only those permitted by law to have access to restricted student information, and 3) there are no additional costs assessed to the student for the measures we use to verify student identity.

## H. Adequacy of faculty resources

See Appendix C for a representative list of faculty who will teach in the proposed program. The proposed program has approximately 31 highly-qualified faculty members. Each is a distinguished and experienced professional, and most have doctorates in their field of expertise. Each has demonstrated a strong commitment to excellence in teaching. Most are practicing computer, network, and system security professionals in local corporations or government, and many hold influential positions in their organizations. Additional adjunct faculty will be added as necessary to teach courses in the proposed degree program. While a majority of the faculty are part-time instructors, the strength of the program depends on the fact that they are full-time practitioners in their disciplines.

## I. Adequacy of library resources

Students have full and complete access to the Milton S. Eisenhower Library on the Homewood campus, which is ranked as one of the nation's foremost facilities for research and scholarship. Its collection of more than three million bound volumes, several million microfilms, and more than 13,000 journal subscriptions has been assembled to support the academic efforts of the University. The interlibrary loan department makes the research collection of the nation available to faculty and students. The library also provides easy access to a wide selection of electronic information resources, including the library's online catalog, and numerous electronic abstracting and indexing tools. Many of the databases are accessible remotely. Librarians help students electronically and the library maintains an extensive web site to take visitors through all of its services and materials. To this are added more than 10,000 audiovisual titles available for on-site consultation.

## J. Adequacy of physical facilities, infrastructure and instructional equipment

All courses will be offered online or at the existing Engineering for Professionals' classroom centers at the Applied Physics Laboratory, the Montgomery County Center, or the Dorsey Center. No new facilities will be needed. The existing EP facilities are reasonably available to provide adequate access to space and equipment.

The online courses do not impact the JHU Whiting School of Engineering's physical facilities and infrastructure. They will be delivered via the JHU-EP online offering infrastructure, which includes the Blackboard course management system and the Adobe Connect video conferencing system. Blackboard is one of the world's leading providers of e-learning systems for higher education institutions. This software focuses on educational outcomes and provides a highly flexible learning environment for students. Both Blackboard and Adobe Connect are supported by the Whiting School and the University's IT infrastructure. These systems provide password-protected online course sites and community management systems that enable ongoing collaborative exchange and provide convenient channels for synchronous and asynchronous learning. The Whiting School already successfully delivers all of its online and web-enhanced courses and programs using these platforms. The University is also outfitted with suitable technical and professional staff and a 24/7 technical help desk to provide technical assistance to the students taking online courses. All of the student services such as application processes, course registration, bookstore, ID service, and advising are currently provided online as well.

## K. Adequacy of financial resources with documentation

See Appendix D for detailed financial information.

## L. Adequacy of provisions for evaluation of program

Once the Master of Science in Data Science program is launched, the program and courses will be evaluated using student surveys and program committee reviews on a regular basis. For example, feedback regarding the appropriateness of course content will be solicited from students every time a course is offered. The program committee will meet annually to assess

course evaluations and other feedback provided by students, faculty and other stakeholders in the program. The faculty will meet regularly for an internal assessment and evaluation of the curriculum. Based on these data, the program committee will implement changes to the program (in terms of curriculum content, course delivery mechanisms, etc.) as necessary.

**M. Consistency with the State's minority student achievement goals**

Any student meeting the admissions requirements can apply to the Master of Science in Data Science. The program will work to help all accepted students improve their workplace competitiveness and reach their professional goals, an aim consistent with the State's minority student achievement goals.

**N. Relationship to low productivity programs identified by the Commission**

Not applicable.

## Appendix A

## Course Descriptions

### Data Science Courses

**675.795 Capstone Project in Data Science (3)**
This course permits graduate students in data science to work with a faculty mentor to explore a topic in depth or conduct research in selected areas. Requirements for completion include submission of a significant paper or project.
*Prerequisites:* Seven data science graduate courses including two courses numbered 605.7xx or 625.7xx or admission to the advanced certificate for post-master's study. Students must also have permission of a faculty mentor, the student's academic adviser, and the program chair. Note that students cannot get credit for both 675.802 Independent Study in Data Science II and 675.795.

**675.801 Independent Study in Data Science I (3)**
This course permits graduate students in data science to work with a faculty mentor to explore a topic in depth or conduct research in selected areas. Requirements for completion include submission of a significant paper suitable to be submitted for publication.
*Prerequisites:* Seven data science graduate courses including two courses numbered 605.7xx or 625.7xx or admission to the advanced certificate for post-master's study. Students must also have permission of a faculty mentor, the student's academic adviser, and the program chairs.

**675.802 Independent Study in Data Science II (3)**
Students wishing to take a second independent study in data science should sign up for this course.
*Prerequisites:* 605.801 Independent Study in Data Science I and permission of a faculty mentor, the student's academic adviser, and the program chair. Note that students cannot get credit for both 675.795 Capstone Project in Data Science and 675.802.

### Computer Science Courses

**605.101 Introduction to Python (0)**
This is a free, noncredit, non-facilitated, six week online course taken for a P/F grade.

**605.102 Introduction to R (0)**
This is a free, noncredit, non-facilitated, six week online course taken for a P/F grade.

**605.201 Introduction to Programming Using Java (3)**
This course enables students without a background in software development to become proficient programmers who are prepared for a follow-on course in data structures. The Java language will be used to introduce foundations of structured, procedural, and object-oriented programming. Topics include I/O, data types, operators, operands, expressions, conditional statements, iteration, recursion, arrays, functions, parameter passing, and returning values.

Students will also be introduced to classes, objects, object references, inheritance, polymorphism, and exception handling. Additional topics include file I/O, searching, sorting, Java Collections, and an introduction to Applets. Students will complete several programming assignments to develop their problem-solving skills and to gain experience in detecting and correcting software errors. (Not for Graduate credit.)
*Prerequisite:* One year of college mathematics.

**605.202 Data Structures (3)**
This course investigates abstract data types (ADTs), recursion, algorithms for searching and sorting, and basic algorithm analysis. ADTs to be covered include lists, stacks, queues, priority queues, trees, sets, and dictionaries. The emphasis is on the trade-offs associated with implementing alternative data structures for these ADTs. There will be four or five substantial Java programming assignments. (Not for Graduate credit.)
*Prerequisites:* One year of college mathematics. 605.201 Introduction to Programming Using Java or equivalent.

**605.203 Discrete Mathematics (3)**
This course emphasizes the relationships between certain mathematical structures and various topics in computer science. Topics include set theory, graphs and trees, algorithms, propositional calculus, logic and induction, functions, relational algebra, and matrix algebra. (Not for Graduate credit.)
*Prerequisite:* Calculus is recommended.605.421 Foundations of Algorithms (3)
This follow-on course to data structures (e.g., 605.202) provides a survey of computer algorithms, examines fundamental techniques in algorithm design and analysis, and develops problem-solving skills required in all programs of study involving computer science. Topics include advanced data structures (red-black and 2-3-4 trees, union-find), recursion and mathematical induction, algorithm analysis and computational complexity (recurrence relations, big-O notation, NP-completeness), sorting and searching, design paradigms (divide and conquer, greedy heuristic, dynamic programming, amortized analysis), and graph algorithms (depth-first and breadth-first search, connectivity, minimum spanning trees, network flow). Advanced topics are selected from among the following: randomized algorithms, information retrieval, string and pattern matching, and computational geometry.
*Prerequisite:* 605.202 Data Structures or equivalent. 605.203 - Discrete Mathematics or equivalent is recommended.

**605.432 Graph Analytics (3)**
Graphs have shown a steady growth in usage with the development of internet, cyber, and social networks. They provide a flexible data structure that facilitates fusion of disparate data sets, however, the processing of large graphs remains a challenging problem. This course introduces algorithms, architectures, and techniques used to address the problem of large scale graph analytics. We will blend graph analytics theory with hands on development of graph storage and the processing systems that support the analytics. By the end of the course, students will get a flavor of graph processing architectures and applications. The course will start by introducing graphs, their properties, and example applications. Implementation of a graph database on a standalone machine will be covered and students will learn how to build their own graph database. Discussions will cover statistical properties of graphs and common graph analytics methods such as graph partitioning, connected components, cliques, and trusses. Distributed

storage and processing architectures that support graph analytics will be examined, and students will learn how to implement a distributed graph processing system. There will be hands on programming assignments.
*Prerequisite:* 605.203 Discrete Mathematics is recommended.

**605.441 Principles of Database Systems (3)**
This course examines the underlying concepts and theory of database management systems. Topics include database system architectures, data models, query languages, conceptual and logical database design, physical organization, and transaction management. The entity-relationship model and relational model are investigated in detail, object-oriented databases are introduced, and legacy systems based on the network and hierarchical models are briefly described. Mappings from the conceptual level to the logical level, integrity constraints, dependencies, and normalization are studied as a basis for formal design. Theoretical languages such as the relational algebra and the relational calculus are described, and high-level languages such as SQL and QBE are discussed. An overview of file organization and access methods is provided as a basis for discussion of heuristic query optimization techniques. Finally, transaction processing techniques are presented with a specific emphasis on concurrency control and database recovery.

**605.445 Artificial Intelligence (3)**
The incorporation of advanced techniques in reasoning and problem solving into modern, complex systems has become pervasive. Often, these techniques fall within the realm of artificial intelligence. This course focuses on artificial intelligence from an agent perspective, and explores issues of knowledge representation and reasoning. Students will participate in lectures and discussions on various topics, including heuristic and stochastic search, logical and probabilistic reasoning, planning, learning, and perception. Advanced topics will be selected from areas such as robotics, vision, natural language processing, and philosophy of mind. Students will complete problem sets and small software projects to gain hands-on experience with the techniques and issues covered.

**605.446 Natural Language Processing (3)**
This course introduces the fundamental concepts and techniques of natural language processing (NLP). Students will gain an in-depth understanding of the computational properties of natural languages and the commonly used algorithms for processing linguistic information. The course examines NLP models and algorithms using both the traditional symbolic and the more recent statistical approaches. It includes treatment of natural languages at the lexical, syntactic, semantic, and pragmatic levels. The course also covers the development of modern NLP systems using statistical and machine learning techniques.
*Prerequisite:* 605.445 Artificial Intelligence or equivalent

**605.448 Data Science (3)**
This course will cover the core concepts and skills in the emerging field of data science. The data science pipeline will be explored in-depth: problem formulation, the acquisition and cleaning of multi-source data sets, data summarization and exploratory analysis, model building, analysis and evaluation; and the presentation of results. Topics covered will include types of data sources and databases, web scraping and APIs, text parsing and regular expressions, experimental design, summary statistics, data visualizations, supervised (regression, logistic regression, decision trees,

random forests, etc.) and unsupervised (clustering, network analysis), machine learning techniques, model evaluation and testing, and the construction of web applications and reports to present results. Students will gain direct experience in solving the programming and analytical challenges associated with Data Science through short assignments and a larger project.
*Prerequisite:* Programming experience in Python is recommended

**605.462 Data Visualization (3)**
This course explores the underlying theory and practical concepts in creating visual representations of large amounts of data. It covers the core topics in data visualization: data representation, visualization toolkits, scientific visualization, medical visualization, information visualization, flow visualization, and volume rendering techniques. The related topics of applied human perception and advanced display devices are also introduced.
*Prerequisite:* Experience with data collection/analysis in data-intensive fields or background in computer graphics (e.g., 605.467 Computer Graphics ) is recommended.

**605.481 Principles of Enterprise Web Development (3)**
This course examines three major topics in the development of applications for the World Wide Web. The first is web site development using HTML and related standards. The second is the implementation of client-side applications using the Java programming language, including user interface development, asynchronous event handling, multithreaded programming, and network programming. Distributed object protocols via RMI or CORBA and distributed database access via JDBC may also be introduced. The third topic is the design of server-side web applications, for which students will examine the underlying Web protocol (HTTP), the development of client-side interfaces (e.g., via HTML forms), and the implementation of server-side programs (e.g., via Java servlets or traditional CGI).

**605.741 Distributed Database Systems: Cloud Computing and Data Warehouses (3)**
This course investigates the architecture, design, and implementation of massive-scale data systems. The course discusses foundational concepts of distributed database theory including design and architecture, security, integrity, query processing and optimization, transaction management, concurrency control, and fault tolerance. It then applies these concepts to both large-scale data warehouse and cloud computing systems. The course blends theory with practice, with each student developing both distributed database and cloud computing projects.
*Prerequisites:* 605.441 Principles of Database Systems and 605.481 Principles of Enterprise Web Development or equivalent knowledge of Java and HTML. Familiarity with "Big Oh" concepts and notation is recommended.

**605.744 Information Retrieval (3)**
A multibillion-dollar industry has grown to address the problem of finding information. Commercial search engines are based on information retrieval: the efficient storage, organization, and retrieval of text. This course covers both the theory and practice of text retrieval technology. Topics include automatic index construction, formal models of retrieval, Internet search, text classification, multilingual retrieval, question answering, and related topics in NLP and computational linguistics. A practical approach is emphasized and students will complete several programming projects to implement components of a retrieval engine. Students will also give a class presentation based on an independent project or a research topic from the IR literature.

**605.746 Machine Learning (3)**
How can machines improve with experience? How can they discover new knowledge from a variety of data sources? What computational issues must be addressed to succeed? These are questions that are addressed in this course. Topics range from determining appropriate data representation and models for learning, understanding different algorithms for knowledge and model discovery, and using sound theoretical and experimental techniques in assessing performance. Specific approaches covered include statistical techniques (e.g., k-nearest neighbor and Bayesian learning), logical techniques (e.g., decision tree and rule induction), function approximation (e.g., neural networks and kernel methods), and reinforcement learning. The topics are discussed in the context of current machine learning and data mining research. Students will participate in seminar discussions and will complete and present the results of an individual project.
*Prerequisite:* 605.445 Artificial Intelligence is recommended, but not required.

**605.748 Semantic Natural Language Processing (3)**
This course introduces the fundamental concepts underlying knowledge representation, semantics and pragmatics in natural language processing. Students will gain an in-depth understanding of the techniques central to computational semantics and discourse for processing linguistic information. The course examines semantic NLP models and algorithms using both the traditional symbolic and the more recent statistical approaches. The course also covers the development of modern NLP systems capable of carrying out dialogue and conversation. This course and 605.446 Natural Language Processing can be taken independently of each other.
*Prerequisite*: 605.445 Artificial Intelligence or equivalent

**605.788 Big Data Processing using Hadoop (3)**
Organizations today are generating massive amounts of data that are too large and too unwieldy to fit in relational databases. So, organizations and enterprises are turning to massively parallel computing solutions such as Hadoop for help. The Apache Hadoop platform, with Hadoop Distributed File System (HDFS) and MapReduce (M/R) framework at its core, allows for distributed processing of large data sets across clusters of computers using the Map and Reduce programming model. It is designed to scale up from a single server to thousands of machines, offering local computation and storage. The Hadoop ecosystem is sizable in nature and includes many sub-projects such as Hive and Pig for big data analytics, HBase for real time access to big data, Zookeeper for distributed transaction process management and Oozie for workflow. This course breaks down the walls of complexity of big data distributed processing by providing a practical approach to developing applications on top of the Hadoop platform. On completing this course, the students would gain an in-depth understanding of how MapReduce and Distributed File Systems work. In addition, they would be able to author Hadoop based Mapreduce applications in Java, and also leverage Hadoop sub-projects to build powerful data processing applications.
*Prerequisite:* 605.481 Principles of Enterprise Web Development or equivalent Java experience.
*Course Notes:* This course may be counted toward a three course track in Databases and Knowledge Management.Applied and Computational Mathematics Courses

**625.201 General Applied Mathematics (3)**
This course is designed for students whose prior background does not fully satisfy the mathematics requirements for admission and/or for students who wish to take a refresher course

in applied mathematics. The course provides a review of differential and integral calculus in one or more variables. It covers elementary linear algebra and differential equations, including first- and second-order linear differential equations. Basic concepts of matrix theory are discussed (e.g., matrix multiplication, inversion, and eigenvalues/eigenvectors). (Not for Graduate credit.)
*Prerequisite:* Two semesters of calculus.
*Course Notes:* This course alone does not fulfill the mathematics requirements for admission to the applied and computational mathematics or the data science programs. Additional course work is required.

**625.250 Applied Mathematics I (3)**
This course covers the fundamental mathematical tools required in applied physics and engineering. The goal is to present students with the mathematical techniques used in engineering and scientific analysis and to demonstrate these techniques by the solution of relevant problems in various disciplines. Areas include vector analysis, linear algebra, matrix theory, and complex variables. (Not for Graduate credit.)
*Prerequisite:* Differential and integral calculus.

**625.251 Applied Mathematics II (3)**
This course is a companion to 625.250. Topics include ordinary differential equations, Fourier series and integrals, the Laplace transformation, Bessel functions and Legendre polynomials, and an introduction to partial differential equations.(Not for Graduate credit.)
*Prerequisite:* Differential and integral calculus. Students with no experience in linear algebra may find it helpful to take 625.250 Applied Mathematics I first.

**625.403 Statistical Methods and Data Analysis (3)**
This course introduces commonly used statistical methods. The intent of this course is to provide an understanding of statistical techniques and guidance on the appropriate use of methodologies. The course covers the mathematical foundations of common methods as an aid toward understanding both the types of applications that are appropriate and the limits of the methods. MATLAB and statistical software are used so students can apply statistical methodology to practical problems in the workplace. Topics include the basic laws of probability and descriptive statistics, conditional probability, random variables, expectation and variance, discrete and continuous probability models, bivariate distributions and covariance, sampling distributions, hypothesis testing, method of moments and maximum likelihood point (MLE) estimation, confidence intervals, contingency tables, analysis of variance (ANOVA), and linear regression modeling.
*Prerequisite:* Multivariate calculus.

**625.415 Nonlinear Optimization (3)**
Although a number of mathematical programming problems can be formulated and solved using techniques from linear and integer problems, there are a wide variety of problems that require the inclusion of nonlinearities if they are to be properly modeled. This course presents theory and algorithms for solving nonlinear optimization problems. Theoretical topics treated include basic convex analysis, first and second order optimality conditions, KKT conditions, constraint qualification, and duality theory. We will investigate an array of algorithms for both constrained and unconstrained optimization. These algorithms include the Nelder-Mead (nonlinear simplex method), steepest descent, Newton methods, conjugate direction methods, penalty methods, and

barrier methods. In parallel with our theoretical and algorithmic development we will consider how to formulate mathematical programs for an assortment of applications including facility location, regression analysis, financial evaluation, and policy analysis. If time permits, we will also address algorithms for special classes of nonlinear optimization problems (e.g., separable programs, convex programs, and quadratic programs).
*Prerequisites & Notes:* Multivariate calculus, linear algebra. Some real analysis would be good but is not required; 625.414 (Linear Optimization) is not required.

**625.461 Statistical Models and Regression (3)**
Introduction to regression and linear and nonlinear models including least squares estimation, maximum likelihood estimation, the Gauss-Markov Theorem, and the Fundamental Theorem of Least Squares. Topics include estimation, hypothesis testing, simultaneous inference, model diagnostics, transformations, multicollinearity, influence, model building, and variable selection. Advanced topics include nonlinear regression, robust regression, and generalized linear models including logistic and Poisson regression.
*Prerequisites:* One semester of statistics (such as 625.403), multivariate calculus, and linear algebra.

**625.462 Design and Analysis of Experiments (3)**
Statistically designed experiments are the efficient allocation of resources to maximize the amount of information obtained with a minimum expenditure of time and effort. Design of experiments is applicable to both physical experimentation and computer simulation models. This course covers the principles of experimental design, the analysis of variance method, the difference between fixed and random effects and between nested and crossed effects, and the concept of confounded effects. The designs covered include completely random, randomized block, Latin squares, split-plot, factorial, fractional factorial, nested treatments and variance component analysis, response surface, optimal, Latin hypercube, and Taguchi. Any experiment can correctly be analyzed by learning how to construct the applicable design structure diagram (Hasse diagrams).
*Prerequisites & Notes:* Multivariate calculus, linear algebra, and one semester of graduate probability and statistics (e.g., 625.403 Statistical Methods and Data Analysis). Some computer-based homework assignments will be given.

**625.464 Computational Statistics (3)**
Computational Statistics is a branch of mathematical sciences concerned with efficient methods for obtaining numerical solutions to statistically formulated problems. This course will introduce students to a variety of computationally intensive statistical techniques and the role of computation as a tool of discovery. Topics include numerical optimization in statistical inference (expectation-maximization [EM] algorithm, Fisher scoring, etc.), random number generation, Monte Carlo methods, randomization methods, jackknife methods, bootstrap methods, tools for identification of structure in data, estimation of functions (orthogonal polynomials, splines, etc.), and graphical methods. Additional topics may vary. Coursework will include computer assignments.
*Prerequisites:* Multivariate calculus, familiarity with basic matrix algebra, graduate course in probability and statistics (such as 625.403).

**625.490 Computational Complexity and Approximation (3)**
This course will cover the theory of computational complexity, with a focus on popular approximation and optimization problems and algorithms. It begins with important complexity concepts including Turing machines, Karp and Turing reducibility, basic complexity classes, and the theory of NP-completeness. It then discusses the complexity of well-known approximation and optimization algorithms, and introduces approximability properties, with special focus on approximation algorithm and heuristic design. The impact of emerging computing techniques, such as massive parallelism and quantum computing, will also be discussed. The course will specifically target algorithms with practical significance and techniques that can improve performance in real-world implementations.
*Prerequisites:* Introductory probability theory and/or statistics (such as 625.403) and undergraduate-level exposure to algorithms and matrix algebra. Some familiarity with optimization and computing architectures is desirable, but not necessary.

**625.492 Probabilistic Graphical Models (3)**
This course introduces the fundamentals behind the mathematical and logical framework of graphical models. These models are used in many areas of machine learning and arise in numerous challenging and intriguing problems in data analysis, mathematics, and computer science. For example, the "big data" world frequently uses graphical models to solve problems. While the framework introduced in this course will be largely mathematical, we will also present algorithms and connections to problem domains. The course will begin with the fundamentals of probability theory, and will then move into Bayesian networks, undirected graphical models, template-based models, and Gaussian networks. The nature of inference and learning on the graphical structures will be covered, with explorations of complexity, conditioning, clique trees, and optimization. The course will use weekly problem sets and a term project to encourage mastery of the fundamentals of this emerging area. (This course is the same as 605.425 Probabilistic Graph Theory.)
*Prerequisite:* Graduate course in probability and statistics (such as 625.403)

**625.721 Probability and Stochastic Process I (3)**
The course is an introduction to probability theory. Topics include sample space, combinatorial analysis, conditional probability, discrete and continuous distributions, expectation and generating functions, laws of large numbers, and central limit theorem. This course is proof oriented and the primary purpose is to lay the foundation for the second course 625.722 and other specialized courses in probability.
*Prerequisites:* Multivariate calculus and 625.403 Statistical Methods and Data Analysis or equivalent.

**625.722 Probability and Stochastic Process II (3)**
This course is an introduction to the theory of discrete-time stochastic processes. Emphasis in the course is given to Poisson processes, renewal theory, renewal reward process, Markov chains, continuous-time Markov chains, birth and death process, Brownian motion, and random walks.
*Prerequisites:* Differential equations and 625.721 Probability and Stochastic Process I or equivalent

**625.725 Theory of Statistics I (3)**
This course covers mathematical statistics and probability. Topics covered include discrete and continuous probability distributions, expected values, moment-generating functions, sampling theory, convergence concepts, and the central limit theorem. This course is a rigorous treatment of statistics that lays the foundation for 625.726 and other advanced courses in statistics.
*Prerequisites:* Multivariate calculus and 625.403 Statistical Methods and Data Analysis or equivalent.

**625.726 Theory of Statistics II (3)**
This course is the continuation of 625.725. It covers method of moments estimation, maximum likelihood estimation, the Cramér-Rao inequality, sufficiency and completeness of statistics, uniformly minimum variance unbiased estimators, the Neyman-Pearson Lemma, the likelihood ratio test, goodness-of-fit tests, confidence intervals, selected non-parametric methods, and decision theory.
*Prerequisites:* 625.725 Theory of Statistics I or equivalent.

**625.734 Queuing Theory with Applications to Computer Science (3)**
Queues are a ubiquitous part of everyday life; common examples are supermarket checkout stations, help desk call centers, manufacturing assembly lines, wireless communication networks, and multi-tasking computers. Queuing theory provides a rich and useful set of mathematical models for the analysis and design of service process for which there is contention for shared resources. This course explores both theory and application of fundamental and advanced models in this field. Fundamental models include single and multiple server Markov queues, bulk arrival and bulk service processes, and priority queues. Applications emphasize communication networks and computer operations, but may include examples from transportation, manufacturing, and the service industry. Advanced topics may vary. (This course is the same as 605.725 Queuing Theory with Applications.)
*Prerequisites:* Multivariate calculus and a graduate course in probability and statistics such as 625.403.

**625.740 Data Mining (3)**
Data mining is a relatively new term used in the academic and business world, often associated with the development and quantitative analysis of very large databases. Its definition covers a wide spectrum of analytic and information technology topics, such as machine learning, pattern recognition, artificial intelligence, statistical modeling, and efficient database development. This course will review these broad topics, and cover specific analytic and modeling techniques such as data cleaning techniques, principal components, regression, decision trees, neural networks, support vector machines, nearest neighbor, clustering, association rules, generalization error, and the holdout, cross-validation, and bootstrap methods. Mathematics underlying these techniques will be discussed and their application to real-world data will be illustrated. Because use of the computer is extremely important when "mining" large amounts of data, we will make substantial use of data mining software tools to learn the techniques and analyze datasets.
*Prerequisites:* Multivariate calculus, linear algebra, and matrix theory (e.g., 625.409), and a course in probability and statistics (such as 625.403). This course will also assume familiarity with multiple linear regression and basic ability to program.

**625.741 Game Theory (3)**
Game theory is a field of applied mathematics that describes and analyzes interactive decision-making when two or more parties are involved. Since finding a firm mathematical footing in 1928, it has been applied to many fields, including economics, political science, foreign policy, and engineering. This course will serve both as an introduction to and a survey of applications of game theory. Therefore, after covering the mathematical foundational work with some measure of mathematical rigor, we will examine many real-world situations, both historical and current. Topics include two-person/N-person game, cooperative/non-cooperative game, static/dynamic game, combinatorial/strategic/coalitional game, and their respective examples and applications. Further attention will be given to the meaning and the computational complexity of finding of Nash equilibrium. (This course is the same as 605.726 Game Theory)
*Prerequisites:* Multivariate calculus, linear algebra and matrix theory (e.g., 625.409), and a course in probability and statistics (such as 625.403).

**625.743 Stochastic Optimization and Control (3)**
Stochastic optimization plays an increasing role in the analysis and control of modern systems. This course introduces the fundamental issues in stochastic search and optimization with special emphasis on cases where classical deterministic search techniques (steepest descent, Newton-Raphson, linear and nonlinear programming, etc.) do not readily apply. These cases include many important practical problems, which will be briefly discussed throughout the course (e.g., neural network training, nonlinear control, experimental design, simulation- based optimization, sensor configuration, image processing, discrete-event systems, etc.). Both global and local optimization problems will be considered. Techniques such as random search, least mean squares (LMS), stochastic approximation, simulated annealing, evolutionary computation (including genetic algorithms), and machine learning are discussed.
*Prerequisites:* Multivariate calculus, linear algebra, and one semester of graduate probability and statistics (e.g., 625.403 Statistical Methods and Data Analysis). Some computer-based homework assignments will be given. It is recommended that this course only be taken in the last half of a student's degree program.

**625.744 Modeling, Simulation, and Monte Carlo (3)**
Computer simulation and related Monte Carlo methods are widely used in engineering, scientific, and other work. Simulation provides a powerful tool for the analysis of real-world systems when the system is not amenable to traditional analytical approaches. In fact, recent advances in hardware, software, and user interfaces have made simulation a "first line" method of attack for a growing number of problems. Areas where simulation-based approaches have emerged as indispensable include decision aiding, prototype development, performance prediction, scheduling, and computer-based personnel training. This course introduces concepts and statistical techniques that are critical to constructing and analyzing effective simulations, and discusses certain applications for simulation and Monte Carlo methods. Topics include random number generation, simulation-based optimization, model building, bias-variance tradeoff, input selection using experimental design, Markov chain Monte Carlo (MCMC), and numerical integration.
*Prerequisites:* Multivariate calculus, familiarity with basic matrix algebra, graduate course in probability and statistics (such as 625.403). Some computer-based homework assignments will be given. It is recommended that this course only be taken in the last half of a student's degree program.

## Appendix B

## Evidence of Compliance with the Principles of Good Practice (as outlined in COMAR 13B02.03.22C)

### (a) Curriculum and Instruction

**(i) A distance education program shall be established and overseen by qualified faculty.**

This program is derived from two well-established site-based programs that have online offerings as well – the Computer Science and Applied & Computational Mathematics master's degree programs. Many of the faculty teaching in the on-site program also serve as online instructors. Any new instructor recruited to teach online would be required to meet the same qualifications as those teaching in the traditional site-based program.

**(ii) A program's curriculum shall be coherent, cohesive, and comparable in academic rigor to programs offered in traditional instructional formats.**

All of the courses in the online program are offered in the traditional, site-based program. Prior to a course being converted for online delivery, the course must be taught minimally twice in-class. A formal online course development process is used to support the course conversion from in-class to online. The online course development process incorporates the Quality Matters™ research-based set of eight standards for quality online course design to ensure the academic rigor of the online course is comparable or better to the traditionally offered course. We are starting to experiment with straight-to-online developments in the expectation that we will get non-regional faculty to develop courses in the future.

**(iii) A program shall result in learning outcomes appropriate to the rigor and breadth of the program.**

The program learning outcomes for the distance education program are identical to the traditional on-site program. The program learning outcomes are derived from input from professionals within the discipline, the program instructors, program leadership and other program stakeholders.

**(iv) A program shall provide for appropriate real-time or delayed interaction between faculty and students.**

The Master of Science in Data Science will be delivered via Blackboard, JHU's course management system. This platform supports asynchronous interaction between faculty and students. Students and faculty also have the option to participate

in optional 'real-time' interaction through weekly web-conference office hours, supported by Adobe Connect.

(v) Faculty members in appropriate disciplines in collaboration with other institutional personnel shall participate in the design of courses offered through a distance education program.

The program has established a process for identifying the appropriate faculty to design an online course. All of the faculty members are selected based on domain expertise, program-related teaching experience and completion of a required online course development training course.

**(b) Role and Mission**

**(i) A distance education program shall be consistent with the institution's mission.**

Refer to Section A.1 in the main body of the proposal.

**(ii) Review and approval processes shall ensure the appropriateness of the technology being used to meet a program's objectives.**

The development of online courses is supported by JHU-EP's Center for Learning Design and Technology (CLDT) professional staff, which includes instructional designers, instructional support specialists and other supporting staff. Each online course development is assigned an instructional designer. The course instructor(s) consults with the instructional designer during the course design process to determine the most effective learning technologies and strategies needed to meet the course learning objectives. The course design goes through multiple reviews by the instructional designer and program chairs. The program chairs are responsible for making sure the course design meets the program's expectations for online courses and that the course learning objectives reflect what your program expects students to achieve after completing this course. Once the online course launches, the assigned instructional designer continually monitors the courses, and consults with the instructor(s) to make adjustments to the course, if needed. All new online courses participate in a mid-term and end-of-term course evaluation process. The mid-term feedback is used to determine if any mid-point term corrections are needed. The end-of-term feedback is used to assess whether further course refinements are needed prior to the next time the course is offered.

**(c) Faculty Support**

**(i) An institution shall provide for training for faculty who teach with the use of technology in a distance education format, including training in the learning management system and the pedagogy of distance education.**

Faculty development support for the development of online course is provided by the JHU-EP Center for Learning Design and Technology (CLDT) professional staff. Faculty have multiple opportunities to receive training in the learning management system, and pedagogy of online learning – these opportunities are presented at various times throughout the year at events such as fall/spring annual faculty meetings, Brown Bag workshops, webinars, and scheduled training sessions. Once an instructor has been identified to develop an online course, they are given access to a set of web-based resources that cover a broad range of topics on online pedagogy, use of instructional technologies and learning management system tutorials. Throughout the online course development the instructor receives direct support and guidance from their assigned instructional designer on variety of online learning related topics.

**(ii) Principles of best practice for teaching in a distance education format shall be developed and maintained by the faculty.**

The JHU-EP CLDT has created a series of online teaching strategies resources. These resources are based best practices from research and other related sources. All new online course instructors are encouraged to review these resources prior to teaching their first online course. New online instructors also receive one-on-one coaching from instructional designers and peer mentors.

**(iii) An institution shall provide faculty support services specifically related to teaching through a distance education format.**

The JHU-EP CLDT provides a wide range of faculty support services for instructors engaged in online instruction. Instructors have access to multi-media specialists, instructional technologists, instructional designers, a training specialist and other institutional support staff to assist them in their role as online instructors. Some of the services provided include instructional technology training, course design support, learning management system training, course production support (i.e. recording studio), video production, and a faculty support help line and email.

**(d) An institution shall ensure that appropriate learning resources are available to students including appropriate and adequate library services and resources.**

The students will have online access to the Milton S. Eisenhower Library on the Homewood campus, which is ranked as one of the nation's foremost facilities for research and scholarship. Its collection of more than three million bound volumes, several million microfilms, and more than 13,000 journal subscriptions has been assembled to support the

academic efforts of the University. The interlibrary loan department makes the research collection of the nation available to faculty and students. The library also provides easy access to a wide selection of electronic information resources, including the library's online catalog, and numerous electronic abstracting and indexing tools. Many of the databases are accessible remotely. Librarians help students electronically and the library maintains an extensive web site to take visitors through all of its services and materials.

**(e) Students and Student Services**

**(i) A distance education program shall provide students with clear, complete, and timely information on the curriculum, course and degree requirements, nature of faculty/student interaction, assumptions about technology competence and skills, technical equipment requirements, learning management system, availability of academic support services and financial aid resources, and costs and payment policies.**

JHU-EP maintains numerous web-based resources to inform prospective students on the information they may need as an online student. These resources include: EP main website (http://ep.jhu.edu); EP online catalog, which includes detailed programmatic information, academic support services, financial aid, costs, policies, etc. and specific information for online learning (refer to http://catalog.ep.jhu.edu/content.php?catoid=20&navoid=630). As new online students are admitted and enrolled, they receive timely emails with important information to help them prepare to become an online student. These emails include information on how to create their JHU login account for the course management systems, technical requirements, available academic support services and new online student orientation course.

**(ii) Enrolled students shall have reasonable and adequate access to the range of student services to support their distance education activities.**

JHU-EP online students have access to the following academic support services:

- **Academic Advising.** Students are assigned an advisor when accepted. Students work individually with the advisor to develop a course of study that meets the requirements of the program and the career goals of the student. The advisor regularly contacts the students to check on progress and answer questions. Courses that deviate from the program plan and have not been approved by an adviser may not count toward certificate requirements. A degree audit tool is provided so students can verify their selections match requirements for the post-master's certificate.

- **Library Services.** Students have online access to the Milton S. Eisenhower Library on the Homewood campus, ranked as one of the nation's foremost facilities for research and scholarship. The interlibrary loan department allows students access to resources at any other university in the nation. The library also provides easy access to a wide selection of

electronic information resources, including the library's online catalog, and numerous electronic abstracting and indexing tools. Many of the databases are accessible remotely. Librarians are available to assist students remotely and the library maintains an extensive web site to take visitors through all its services and materials.

- **Services with Students with Disabilities.** The Johns Hopkins University is committed to making all academic programs, support services, and facilities accessible to qualified individuals. Students with disabilities who require reasonable accommodations can contact the EP Disability Services Administrator.

- **Johns Hopkins Student Assistance Program.** The Johns Hopkins Student Assistance Program (JHSAP) is a professional counseling service that can assist students with managing problems of daily living. Stress, personal problems, family conflict, and life challenges can affect the academic progress of students. JHSAP focuses on problem-solving through short-term counseling. Accessing the service is a simple matter of a phone call to arrange an appointment with a counselor. Online students may call a phone number for consultation and will be directed to the appropriate resource or office. JHSAP services are completely confidential. The program operates under State and Federal confidentiality legislation and is HIPAA compliant.

- **Transcript Access.** Official transcripts will be mailed upon written request of the student at no charge.

- **Student ID JCard.** The JCard serves as the student's University identification card. This card is mailed to the home address of every registered student. The JCard acts as the university library card, which enables students to check out books from the Homewood Eisenhower Library or at any of the campus center libraries, and provides access to many computer laboratories.

**(iii) Accepted students shall have the background, knowledge, and technical skills needed to undertake a distance education program.**

All accepted online students are required to have met the admission requirements stated for the degree program. New online students are strongly encouraged to complete the "New Online Student Orientation" course prior to beginning their first online course. This course covers a broad range of topics on how to be a successful online student such as: online student learning expectations, how to access the library, how to conduct online research, and how to participate in online discussions.

**(iv.) Advertising, recruiting, and admissions materials shall clearly and accurately represent the program and the services available.**

All relevant program information is kept up to date on the JHU-EP web site (http://ep.jhu.edu).

**(f) Commitment to Support**

**(i) Policies for faculty evaluation shall include appropriate consideration of teaching and scholarly activities related to distance education programs.**

Faculty teaching online courses are strongly encouraged to participate in minimally one to two professional development opportunities annually to improve their online teaching skills.

**(ii) An institution shall demonstrate a commitment to ongoing support, both financial and technical, and to continuation of a program for a period sufficient to enable students to complete a degree or certificate.**

Please see sections J and K of the proposal.

**(g) Evaluation and Assessment**

**(i) An institution shall evaluate a distance education program's educational effectiveness, including assessments of student learning outcomes, student retention, student and faculty satisfaction, and cost-effectiveness.**

Please see section L of the proposal.

**(ii) An institution shall demonstrate an evidence-based approach to best online teaching practices.**

The instructional design and faculty support staff in the JHU-EP Center for Learning Design & Technology unit continually participates in professional development activities to keep abreast of evidence-based approaches to online teaching practices. These online teaching practices are then incorporated into the new online instructor training sessions.

**(iii) An institution shall provide for assessment and documentation of student achievement of learning outcomes in a distance education program.**

As part of the online course design process, course assessments are required to be aligned with stated course learning outcomes. The JHU-EP program, where appropriate, incorporates authentic-based learning assessments that demonstrate student's application of learned concepts.

## Appendix C

## Faculty

| First Name | Last Name | Terminal Degree | Field | Academic Title | Status | Course(s) to Teach |
|---|---|---|---|---|---|---|
| Mostafa | Aminzadeh | Ph.D | Statistics | Lecturer | Part-Time | 625.721 Probability and Stochastic Process I<br>625.722 Probability and Stochastic Process II<br>625.725 Theory of Statistics I<br>625.726 Theory of Statistics II |
| Barry | Bodt | Ph.D | Statistics | Lecturer | Part-Time | 625. 403 Statistical Methods and Data Analysis |
| Stephen | Butcher | M.S. | Computer Science | Lecturer | Part-Time | 605.445 Artificial Intelligence<br>605.448 Data Science |
| Beryl | Castello | Ph.D | Applied Mathematics and Statistics | Senior Lecturer | Full-Time | 625.415 Nonlinear Optimization |
| Matthew | Henry | Ph.D. | Systems Engineering | Lecturer | Part-Time | 625.741 Game Theory |
| Eleanor | Chlan | Ph.D. | Computer Science | Lecturer | Part-Time | 605.202 Data Structures<br>605.203 Discrete Mathematics<br>605.462 Data Visualization |
| L. Douglas | Ferguson | M.S.<br>Masters | Electrical Engineering Management | Lecturer | Part-Time | 605.201 Programming in Java |
| Stacy | Hill | Ph.D. | Systems Science and Mathematics | Lecturer | Part-Time | 625.744 Modeling, Simulation, and Monte Carlo |
| Hsien-Ming | Hung | Ph.D. | Statistics | Lecturer | Part-Time | 625.461 Statistical Models and Regression |

| | | | | | | |
|---|---|---|---|---|---|---|
| Charles | Kann | D.Sc. | Computer Science | Lecturer | Part-Time | 605.202 Data Structures |
| Satyendra | Kumar | Ph.D. | Biomedical Engineering | Lecturer | Part-Time | 605.446 Natural Language Processing<br>605.748 Semantic Natural Language Processing |
| Dar-Ning | Kung | Ph.D. | Mechanical Engineering | Lecturer | Part-Time | 605.441 Principles of Database Systems |
| William | Lew | Ph.D. | Mathematical Statistics | Lecturer | Part-Time | 605.421 Foundations of Algorithms |
| Simon | Liu | Ph.D. | Computer Science | Lecturer | Part-Time | 605.441 Principles of Database Systems |
| Dima | May | M.S. | Computer Science | Lecturer | Part-Time | 605.788 Big Data Processing Using Hadoop |
| Paul | McNamee | Ph.D. | Computer Science | Lecturer | Part-Time | 605.744 Information Retrieval |
| Aaron | Navarro | Ph.D. | Computer-Based Instruction | Lecturer | Part-Time | 605.744 Information Retrieval |
| Christine | Nickel | Ph.D. | Applied Mathematics and Statistics | Lecturer | Part-Time | 625.464 Computational Statistics<br>625.734 Queuing Theory with Applications to Computer Science |
| Lixin | Qie | Ph.D. | Biochemistry | Lecturer | Part-Time | 605.201 Programming in Java |
| Cheryl | Resch | M.S<br>M.S. | Computer Science,<br>Mechanical Engineering | Lecturer | Part-Time | 605.202 Data Structures |
| Benjamin | Rodriguez | Ph.D. | Electrical and Computer Engineering | Lecturer | Part-Time | 605.421 Foundations of Algorithms |
| John | Sadowsky | Ph.D. | Mathematics | Lecturer | Part-Time | 605.421 Foundations of Algorithms |
| Cetin | Savkli | Ph.D. | Theoretical Physics | Lecturer | Part-Time | 605.432 Graph Analytics |

## Appendix D

## Finance Information

| TABLE 1: RESOURCES | 2016 | 2017 | 2018 | 2019 | 2020 |
|---|---|---|---|---|---|
| 1. Reallocated Funds | $0 | $0 | $0 | $0 | $0 |
| 2. Tuition/Fee Revenue (c + g below) | $105,930 | $221,394 | $404,874 | $604,419 | $884,264 |
| a. Number of F/T Students | 0 | 0 | 0 | 0 | 0 |
| b. Annual Tuition/Fee Rate | - | - | - | - | - |
| c. Total F/T Revenue (a x b) | $0 | $0 | $0 | $0 | $0 |
| d. Number of P/T Students | 10 | 20 | 35 | 50 | 70 |
| e. Credit Hour Rate | $1,177 | $1,230 | $1,285 | $1,343 | $1,404 |
| f. Annual Credit Hour Rate | 9 | 9 | 9 | 9 | 9 |
| g. Total P/T Revenue (d x e x f) | $105,930 | $221,394 | $404,874 | $604,419 | $884,264 |
| 3. Grants, Contracts & Other Ext Sources | $0 | $0 | $0 | $0 | $0 |
| 4. Other Sources | $0 | $0 | $0 | $0 | $0 |
| **TOTAL (Add 1 – 4)** | **$105,930** | **$221,394** | **$404,874** | **$604,419** | **$884,264** |

**Resources Narrative**

1. Reallocated Funds: The proposed program will be funded by tuition revenue, and will make no use of reallocated funds.
2. Tuition and Fee Revenue: Revenue is based on projected enrollments for the program. The enrollment projections in Table 1 are a reasonable estimate based on other new programs in the Whiting School of Engineering and reflect the existing and growing demand for graduates of professional engineering and science master's degree programs at other universities in the nation. The Master of Science in Data Science is a part-time degree program, so no full-time students are expected.
3. Grants and Contracts: No grants or contacts are required for the successful implementation of the program.
4. Other Sources: The program does not expect any funding from other sources.

**NOTE**: The resources and expenditures data for the Master of Science in Data Science are combined with those for the Post-Master's Certificate in Data Science, as they share the same courses, and all resources and expenditures in these programs are course-based.

| TABLE 2: EXPENDITURES | 2016 | 2017 | 2018 | 2019 | 2020 |
|---|---|---|---|---|---|
| 1. Faculty (b + c below) | $27,524 | $56,149 | $76,362 | $107,098 | $148,963 |
| a. # Sections offered | 1 | 4 | 7 | 10 | 14 |
| b. Total Salary | $25,485 | $51,989 | $70,706 | $99,165 | $137,929 |
| c. Total Benefits | $2,039 | $4,159 | $5,656 | $7,933 | $11,034 |
| 2. Admin. Staff (b + c below) | $41,606 | $42,438 | $43,287 | $69,733 | $71,128 |
| a. # FTE | 0.4 | 0.4 | 0.4 | 0.7 | 0.7 |
| b. Total Salary | $23,616 | $24,089 | $24,571 | $40,838 | $41,655 |
| c. Total Benefits | $17,989 | $18,349 | $18,716 | $28,895 | $29,473 |
| 3. Support Staff (b+c below) | $19,112 | $19,399 | $9,845 | $9,993 | $10,143 |
| a. # FTE | 0.25 | 0.25 | 0.125 | 0.125 | 0.125 |
| b. Total Salary | $14,210 | $14,423 | $7,320 | $7,430 | $7,541 |
| c. Total Benefits | $4,902 | $4,976 | $2,525 | $2,563 | $2,602 |
| 4. Equipment | $0 | $0 | $0 | $0 | $0 |
| 5. Library | $0 | $0 | $0 | $0 | $0 |
| 6. New or Renovated Space | $0 | $0 | $0 | $0 | $0 |
| 7. Other Expenses | $9,631 | $59,516 | $107,278 | $157,853 | $227,623 |
| **TOTAL (Add 1 – 7)** | **$97,872** | **$177,502** | **$236,772** | **$344,677** | **$457,858** |

**Expenditures Narrative**

1. Faculty: The Engineering for Professionals lecturers are paid $8,495 (for FY15) per course taught or developed. For years 2 – 5 an additional 2% was added to the salary rate. The fringe rate is estimated at 8%.
2. Administrative Staff: Includes pro-rated salaries for the Program Manager and Program Chairs.
3. Support Staff: Includes pro-rated salaries for F/T Instructional Designers to assist in developing online courses.
4. Equipment: No direct equipment costs are identified.
5. Library: Existing library facilities are sufficient to meet the needs of the program.
6. New or Renovated Space: No new or renovated space will be needed.
7. Other Expenses: Indirect program costs (per enrollment) are provided here.