**New Degree Program Proposal**
**Master of Science in Data Science and Strategic Analytics**

**(This new program will replace the current Master of Science in Computational Science)**

**Proposal Faculty**

Preparation
J. Russell Manson, Ph.D., P.E., Professor of Physics
[Director of Master's Program in Computational Science]

Janet Wagner, Ph.D., Dean, School of Business
[Adjunct Professor in Operations Research, School of Business]

**P m [(P m [(P m [(P m7) P1(im2( o1.C >></MCID Professor of Computer Science and Information Systems 0 ></MCID 31.72 Tm**

Adeline Koh, Ph.D. Associate Professor, School of Arts and Humanities
[Director of Digital Humanities at Stockton]
Jessica Fleck, Ph.D. Associate Professor of Psychology

Clifton Baldwin, Ph.D. Senior systems engineer, Federal Aviation Administration

**Additional Support:**

Lewis Leitner, Ph.D., Dean, Graduate and Continuing S-2(iT M6nt)- di(e)4( a)s

**Introduction**

***"Information is the oil of the 21st century, and analytics is the combustion engine." - Peter Sondergaard, Senior Vice President, Gartner Research.***

In 2013, IBM estimated that two and a half million terabytes of data are created every day. For the layman, this is the equivalent of over 300 million high definition movies. The National Security Agency (NSA) gathers as much information as is stored in the Library of Congress every 6 hours.[1] Ninety percent of the world's data was generated in just the past two years. Data is created by, among others:

1. Individuals (through social networks and smartphones)

2. Machines (through real-time, network connected sensors – "the Internet of things")

3. Business and commerce (e.g. transaction records and financial data)

4. Higher Education (e.g. registrations, completion and retention data, faculty workload)

5. Science (e.g. bioinformatics, large scale simulation, personal health records)

Making sense of this vast sea of data for the use and benefit of society is considered

**Business:** Using historical discounting data from a department chain store at two hundred locations to predict how sales vary with department.

**Entertainment:** Conducting a sentiment analysis on the tweets about summer blockbuster movies and using the data to predict future box office receipts.

**History:** Developing a geospatial database of conflicts occurring during the Scottish Wars of Independence (1296-1328).

**Health:** Predicting disease likelihood by exploring and correlating patient case history and genetic databases.

**Criminal Justice**: Gathering and visualizing real time crime statistics for a city for the sake of efficient resource deployment.

**Education**: Creating a web-based dashboard for describing student performance metrics across a school district.

**Science:** Analyzing the jpg images of one million galaxies to categorize them according to their morphology.

**Sustainability:** Creating a "smart building": a building fitted with sensors collecting data on all aspects of the building's efficiency thus allowing real time adjustments to save energy.

We propose this new degree in response to this exploding "big data" industry sector. Stockton University is extremely well-positioned to deliver such a degree because of existing resources - including faculty (full-time and adjunct), computer equipment, curriculum and courses - within existing schools and programs.

Specifically, we wish to create a Master of Science in Data Science and Strategic Analytics (MS DSSA) by pooling and re-structuring present Stockton resources (faculty, equipment, courses) to better reflect the emerging data science industry. Students in the existing computational science master's degree program are already being introduced to data science courses. Program faculty felt we had to offer such courses to properly serve our students. Students in the existing degree would be given the choice as to whether they want to complete the MSCP degree or switch to pursuing a MS

## 1. Program Objectives

**2. Evaluation and Learning Outcomes Assessment Planning**

The program will conduct learning outcomes assessment on a scheduled basis and report its findings on an annual basis (Graduate Program Director Annual Report and another mechanism to reach the larger Stockton community). We will collect reliable direct and indirect evidence of student learning that will bear upon making improvements in the curriculum, teaching methods and learning outcomes. Overtime, programmatic changes will be based on learning assessment results.

The proposed Master of Science in Data Science and Strategic Analytics will align its curriculum so that it will be consistent with the broad objectives/outcomes in the following Stockton University Essential Learning Outcomes (taken to the graduate level):

” Program Competency…..(integrating concepts, theories and key principles)
” Quantitative reasoning….(working comfortably with numbers and math concepts)
” Critical Thinking………..(formulating effective, balanced perspectives issues and to6(s 11.04 5(u)-4(

At the end of this program, students will be able to ……

| | | |
|---|---|---|
| | Total = 280 Category = 139 | One classes per semester--online--19% |
| | | One classes per semester--hybrid--18% |
| | | One classes per semester--F to F--13% |

Total = 280
Category
=151

*B. Equipment, Materials, Library*

Based on the nature of the program, we anticipate no additional physical resources beyond the current Stockton Campus and Off-Campus facilities. We plan to utilize the resources and expertise offered by the Office of E-Learning and the Office of Computer Services. These offices will continue to support the online components of this program as they have already being doing for the MSCP. Blackboard will be utilized extensively. The data science industry makes extensive use of open license and open source materials; the largest user base is in Linux, Python and R which are all free to use. Our courses will all be based on open license software.

In terms of library support, the college library has significant holdings in computer and computational studies. Data science is an emerging discipline and so it draws from a variety of existing resources. These include both print and internet sources. In addition, faculty will use to online journals and reviews, as well as blogs. We do not anticipate significant additional resources to be needed either in equipment or software to support this program.

*C. Accreditation*

Regional accreditation will continue as per the Middle States Association Commission on Higher Education. No specialized accreditation is currently required or will be sought for this program, however this may change as the industry/discipline matures.

bound by the academic progress requirements of the College's graduate school, covered in the Graduate *Bulletin*.

The program will work closely with the office of e-Learning and the faculty for this program will work with a common online structure. All 10 courses will have a similar online structure and/or "look and feel" to make the student experience an "integrated" experience, instead of ten courses that are merely "co-located".

**Table 1: Curriculum and ideal sequence of courses (accelerated 7-week course structure model)**

| Course | Description | Sequence |
|---|---|---|
| **REQUIRED CORE COURSES [These five courses also constitute the certificate]** | | |
| **DSSA 5001 Introduction to data science and analytics** | A survey course introducing the student to the field of b data. Various application areas will be studied along. B introduction to data gathering, analysis and visualizatio Introduction to tools and technologies. Overview of cybersecurity. | Fall: subterm A |
| **DSSA 5101 Data structures and exploration** | Manipulating and cleaning datasets; storing data in appropriate data structures. Techniques for exploring la datasets introduced. Extensive use of industry standard software tools and operating systems/scripting languag | Fall: subterm A |
| **DSSA 5102 Data gathering and warehousing** | Gathering data from public and private databases. Data and web mining. Structured and unstructured database usage. Extensive use of industry standard software too and operating systems/scripting languages. Data secur | Fall: subterm B |
| **DSSA 5103 Data Visualization** | Interactive visualization of large datasets. Visualization delivery by web interface. Emphasis on integrity, parsimony and aesthetics of data representation. Exten use of industry standard software tools and operating systems/scripting languages. | Fall: subterm B |