

FLORIDA ATLANTIC UNIVERSITY™

Graduate Programs—New CERTIFICATE PROGRAM

UGPC APPROVAL _____

UFS APPROVAL _____

CATALOG _____

 DEPARTMENT: COMPUTER/ELECTRICAL ENGINEERING AND
COMPUTER SCIENCE

COLLEGE: ENGINEERING AND COMPUTER SCIENCE

PROGRAM NAME:

Certificate Program in Big Data Analytics
EFFECTIVE DATE

(PROVIDE TERM/YEAR)

FALL 2015

PLEASE EXPLAIN THE REQUESTED CHANGE(S) AND OFFER RATIONALE BELOW AND/OR ATTACHED:

THE DEPARTMENT OF COMPUTER AND ELECTRICAL ENGINEERING AND COMPUTER SCIENCE (CEECS) IS PROPOSING A CERTIFICATE PROGRAM IN BIG DATA ANALYTICS. ALL THE COURSES ARE OFFERED BY THE CEECS DEPARTMENT. A DESCRIPTION OF THE PROPOSED CERTIFICATE PROGRAM IS ATTACHED. THE CERTIFICATE PROGRAM HAS BEEN APPROVED BY THE DEPARTMENT GRADUATE COMMITTEE.

Faculty contact, email and complete phone number:

 Frederick Bloetscher, Ph.D., P.E.
239-250-2423

Consult and list departments that might be affected by the change and attach comments.

None – the change is only to this department based on classes offered

Approved by:

Department Chair:



College Curriculum Chair:

College Dean:

UGPC Chair:

Graduate College Dean:

UFS President:

Provost:

Date:

05/21/2015

8/3/15

8/11/2015

Email this form and syllabus to UGPC@fau.edu one week before the University Graduate Programs Committee meeting so that materials may be viewed on the UGPC website prior to the meeting.

**Department of Computer and Electrical Engineering and Computer Science,
Florida Atlantic University
Online Certificate Program in Big Data Analytics to be offered through DEDECS**

Proposal: A Distance Learning Certificate Program in Big Data Analytics

Introduction and Rationale: The digital age is here to stay. As a result, organizations now own and have access to unfathomable amounts of data. The potential of this data is largely untapped; new technologies and efforts are needed to move on to the next phase of the digital revolution- the *data* revolution. Hardware and software technologies have been advancing since the inception of electronic computing to bring us to where we are today. Enormous amounts of data are captured and stored at lighting speed, whether it be through software products such as social media and smart phone apps, various hardware sensors tied to countless entities, health data from doctors' offices and physiological measurements, or government data from census and other reporting programs.

The collective problem of dealing with and extracting information from these large sources of data has been coined Big Data. While this term is an often-overused buzzword, a concrete definition of Big Data is elusive and highly debated. The first sighting of the term was in a 1997 paper written by NASA, quantifying Big Data as data too large to fit into computer memory or a hard drive. The problem with this definition is that as computing resources and data management processes mature, a single measurement of data size cannot effectively differentiate "big" from "small" data. The relative nature of the term warrants a more operational and multi-dimensional approach to quantifying Big Data. Gartner research defines Big Data as "high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making." Volume refers to the sheer amount of data, multi-terabytes to petabytes of which can be stored in data warehouses. Velocity refers to the speed at which new data is created; traditional analytical methods cannot keep up with the data coming in. Variety refers to the complexity of the data, as data can come from many sources and have many different formats. The last part of Gartner's definition is key for Big Data analysis: data is useless until it can concretely contribute to a process or industry.

Data scientists, those skilled in handling and gaining insights from Big Data, are in high demand across various industries, but there are simply not enough trained professionals to satisfy the world's data needs. The Wall Street Journal estimates that 80% of U.S. data science jobs created between 2010 and 2011 have not yet been filled. Several universities in the State of Florida offer Master's degrees and/or certificates in business intelligence, or domain-specific informatics (such as Healthcare Informatics). The University of South Florida offers a graduate program in Analytics and Business Intelligence. A couple of examples of Big Data-related programs outside Florida include Stanford's Mining Massive Data Sets Graduate Certificate and University of Maryland University College's degrees and certificates in data analytics. Several Massive Open Online Courses (MOOCs) exist to fill the need for data scientists, such as Udacity's Data Analyst

Nanodegree, and Coursera's (taught by Johns Hopkins University) Data Science Specialization.

The Department of Computer and Electrical Engineering and Computer Science is uniquely positioned to offer a distinctive certificate program in Big Data Analytics. The NSF Big Data Training & Research Laboratory offers students access to a large Big Data computing cluster that would be difficult to find and/or utilize at other universities. While other Florida universities have BI/Data analytics graduate programs, this certificate will be the first in the area of "Big Data". Additionally, the College of Engineering and Computer Science has relationships with such companies as LexisNexis and Modernizing Medicine that have real Big Data Analytics needs and will support this certificate.

It is expected that about 10 students would enroll in the program in the first year; the department's goal is to reach an enrollment number of 30 for the online certificate program by the year 2017. All of the courses required for the new certificate program are already offered online through DEDECS (<http://www.dedecs.fau.edu/>). It should be noted that the present MS (non-thesis) programs in both Computer Science and Computer Engineering require 33 credits. The students enrolling for the online program will be advised by the members of the department's graduate committee and the graduate program coordinator.

Admission and Graduation Requirements: The certificate program will be open to students who have a BS degree in computer science and a related field of engineering, a GPA of at least 3.0 or equivalent (to ensure equivalency to graduate standing) and must satisfy the prerequisites required for each course in the program. Four courses in the program must be completed with a GPA of 3.0 or better. All course materials will be in English and all international students must demonstrate proficiency in English to enter the program.

Curriculum: The four courses required for the certificate program must be selected from the following eight courses (all 3 credit hour courses).

- CAP6673 Data Mining and Machine Learning
- CAP6777 Web Mining
- CAP6778 Advanced Data Mining and Machine Learning
- CAP6776 Information Retrieval
- CAP6771 Data Mining for Bioinformatics
- CAP6688 Social Networks and Big Data Analytics
- CAP6780 Big Data Analytics with Hadoop
- CAP5615 Intro to Neural Networks