SECTION: 13
DATE:
April 21, 2017

BOARD OF REGENTS

EASTERN MICHIGAN UNIVERSITY

RECOMMENDATION

NEW ACADEMIC PROGRAM

ACTION REQUESTED

It is recommended that the Board of Regents approve a New Academic Program: *Data Science and Analytics Major (Bachelor of Science).*

SUMMARY

Analyzing large data sets—so-called big data—has become an essential basis of both science and business, underpinning new waves of productivity growth, innovation, and consumer behavior analytics. To remain competitive and maximize efficiencies in the coming decade companies need to invest in big data technologies. Equally important will be the investment in data scientists' proficiency in advanced statistics and machine learning. Nobody doubts that companies will continue to ramp up hiring, recruitment, and training of data scientists. But there seems to be a growing alarm that we won't have enough data scientists to go around.

The *Data Science and Analytics Major* is an interdisciplinary program designed to provide broad knowledge and the technical competency required to manage and analyze the ever-growing body of data in the industry, government, science, and nonprofit organizations. Graduates will be able to organize, aggregate and analyze rapidly generated data from diverse sources. Such skills are already in high demand in such diverse disciplines as econometrics, bioinformatics, remote sensor networks, and health care analytics. Demand for highly qualified data scientists in these fields will only grow in the future, providing opportunities for our graduates to contribute to scientific research, and engage with the community at the local and global level. The program is both practical and theoretical, providing students the skills to contribute immediately upon graduation and the ability to assimilate emerging technologies as they develop.

FISCAL IMPLICATIONS

Program costs will be absorbed by the current Academic Affairs budget.

ADMINISTRATIVE RECOMMENDATION

The proposed Board action has been reviewed and is recommended for Board approval.

[redacted]

University Executive Officer

3-30-17

Date

EASTERN MICHIGAN UNIVERSITY
DIVISION OF ACADEMIC AND STUDENT AFFAIRS
INTEROFFICE MEMORANDUM

TO: Chris Shell, Registrar
Augustine Ikeji, Department Head, Computer Science
Debra Ingram, Department Head, Mathematics
David Chou, Department Head, Computer Information Systems

FROM: Michael Tew, Interim Director, Undergraduate Studies [redacted]

SUBJECT: **Data Science and Analytics Major (New Program)**

DATE: February 15, 2017

The attached proposal from the Computer Science Department, Mathematics Department, Computer Information Systems Department, College of Arts and Sciences, and the College of Business for a new undergraduate program in **Data Science and Analytics Major**, is approved.

The effective date will be determined following consideration by the Academic Officers Committee, Michigan Association of State Universities and the Eastern Michigan University Board of Regents.

If you have any questions, please contact Evan Finley, Course and Program Development Associate (487-8954, efinley2@emich.edu).

cc: Rhonda Longworth, Interim Provost
Winifred Martin, Executive Assistant, Provost Office
Anne L. Balazs, Interim Dean, College of Business
Toni Jones, Interim Associate Dean, College of Business
Kathleen Stacey, Interim Dean, College of Arts and Sciences
Kate Mehuron, Associate Dean, College of Arts and Sciences
Faculty Senate
Calvin McFarland, Director, University Advising and Career Development
Pat Cygnar, Director, Community College Relations
Colleen Kibin, Assistant Director, Community College Relations
Sarah Kersey Otto, Director, Career Development & Outreach
Extended Programs
John Feldkamp, Assistant Director, Honors College
Ramona Milligan, Coordinator, Registration
Carol Evans, Transfer Equivalency Coordinator, Records & Registration
Erin Burdis, Assistant Registrar
Karen Schiferl, Director, Student-Athlete Support Services
Mary Butkovich, Halle Library
Ann Richards, Assistant Director, Admissions Processing
Bin Ning, Assistant Vice President and Executive Director, IRIM
Andrew Ross, Mathematics Department
Steven Mrdalj, Computer Information Systems
William Sverdlik, Computer Science
Original, Course and Program Development

Eastern Michigan University

2017-2018 Undergraduate Catalog (Working Draft)

Data Science and Analytics | BS

New Program | *effective date TBD*

The effective date will be determined following consideration by the Academic Officers Committee, Michigan Association of State Universities and the Eastern Michigan University Board of Regents.

As you pursue a **Bachelor of Science** in **Data Science and Analytics**, you will learn the interdisciplinary mix of data handling, analysis, and business skills you need to help manage and use the ever-growing body of data available in industry, government, science, and nonprofit organizations.

Learn

As a DSA major, you will gain expertise in computer science/programming, statistics, and business analysis. You will be able to organize, aggregate, store, and use rapidly generated data from diverse sources—data sets so large they need to be stored and processed in a distributed manner rather than on just one computer. You will learn algorithms to discover knowledge and make decisions from extremely high-dimensional data sets. Concentration options include Data Science Theory, Web Analytics, or Advanced Data Analysis.

Opportunities

Students who graduate from the DSA program will be able to enter a career immediately if they choose, or continue on to advanced degrees. Opportunities include any sector of business or industry, research institutions, regulatory agencies, nonprofit organizations, and consulting firms.

As they study, students have the opportunity to participate in department seminars in the Mathematics, Computer Science, and Computer Information Systems departments. Students will also participate in a capstone course with a large, real-world project, and are encouraged to present their work at the Undergraduate Symposium and professional conferences.

Program Information

Data Science and Analytics Program, *College of Arts & Sciences*

TBD Ph.D. | *Program Administrator*

Program Office | TBD | 734.487.XXXX | XXXXX@emich.edu

Advisor Information

Andrew Ross, Ph.D. | *Mathematics* | 515 Pray-Harrold | 734.487.1444 | andrew.ross@emich.edu

Yaman Roumani, Ph.D. | *Computer Information Systems* | 412 Owen | 734.487.2454 | yaman.roumani@emich.edu

William Sverdlik. Ph.D. | *Computer Science* | 512E Pray-Harrold | 734.487.7081 | wsverdlik@emich.edu

General Education Requirements:

For specific General Education requirements, click here (or print a General Education Worksheet)

Major Requirements: 70-74 hours

Required Courses: 58-59 hours

- COSC 146 - Applied Programming and Scripting 3 hrs
- COSC 111 - Introduction to Programming 3 hrs
- COSC 112 - Introduction to Programming Online Lab 1 hr
- COSC 211 - Programming Data Structures 3 hrs
- COSC 212 - Programming Data Structures Online Lab 1 hr
- COSC 311 - Algorithms and Data Structures 4 hrs
- COSC 472 - Big Data I 4 hrs
- COSC 481W - Software Engineering and Senior Project | GEWI 4 hrs
- IS 385 - Data Warehousing 3 hrs
- IS 410 - Data Mining 3 hrs
- MATH 120 - Calculus I | GEQR 4 hrs
- MATH 121 - Calculus II 4 hrs
- MATH 122 - Elementary Linear Algebra 3 hrs

  One course from the following
- DS 265 - Business Statistics I 3 hrs
- STAT 360 - Statistical Methods 3 hrs

  One course from the following
- DS 465 - Applied Linear Statistical Models 3 hrs
- STAT 461 - Linear Regression Analysis 3 hrs

  One course from the following
- COSC 231 - Internet-based Computing 3 hrs
- IS 247 - Introduction to Web Application Development 3 hrs

  One course from the following
- IA 212 - Open Systems Platform and Network Administration 3 hrs
- IS 416 - System Software Administration 3 hrs

One course from the following

- COB 200L4 - Introduction to Business I GELB 3 hrs
- IS 350 - Enterprise Resource Planning and Architecture 3 hrs

One course from the following

- COSC 471 - Database Principles 4 hrs
- IS 380 - Data and Information Management 3 hrs

[1]IS 215 prerequisite will be waived for students in this major.

Concentration Courses: 12-15 hours

Select one of the following concentrations

Theory Concentration: 13-15 hours

Choose four courses from the following

- COSC 221 - Computer Organization I 3 hrs
- COSC 423 - Computer Operating Systems 4 hrs
- COSC 439 - Computing Network Principles 4 hrs
- MATH 223 - Multivariable Calculus 4 hrs
- MATH 419W - Introduction to Stochastic Mathematical Modeling I GEWI 3 hrs
- STAT 370 - Probability and Statistics I 3 hrs

Web Analytics Concentration: 12 hours

- IS 310 - Social Media for Business 3 hrs
- IS 405 - Web Content Mining 3 hrs
- IS 422 - Social Media Network Analysis 3 hrs
- MKTG 339 - Google Analytics and Landing Page Optimization 3 hrs

Advanced Data Analysis Concentration: 12-13 hours

Choose four courses from the following

- COSC 462 - Introduction to Information Retrieval 3 hrs
- COSC 473 - Big Data II 4 hrs
- MATH 419W - Introduction to Stochastic Mathematical Modeling I GEWI 3 hrs
- STAT 462 - Design and Analysis of Experiments 3 hrs
- STAT 468 - Introduction to Biostatistics 3 hrs
- STAT 474W - Applied Statistics I GEWI 3 hrs

Minor Requirement:

This major does not require a minor.

Program Total:

Students must earn a minimum total of 124 credits at the 100-level or above.

Critical Graduation Information

The following are minimum requirements for all bachelor's degrees awarded by Eastern Michigan University. Some majors and minors require more than the minimum in one or more of the areas below; students are urged to consult the on-line catalog for the requirements of their particular programs.

- Earn a minimum total of 124 credits at the ***100-level and above*. Courses with numbers below 100 will not be counted toward this degree requirement. At most 8 credit hours of physical education (PEGN) activity courses will be counted toward this requirement.**
- Meet the requirements of the General Education program (see *information below*).
- Complete a Writing Intensive (GEWI) Course in your major.
- Earn a minimum of 60 credits from a four-year college or university; **courses taken at community colleges cannot be used to meet this requirement.** (Some formal program-to-program articulation agreements modify this requirement. See specific agreements for details.)
- Earn a minimum of 30 credits from courses taken at EMU.
- Complete 10 of the last 30 hours for the degree from courses taken at EMU.
- Have a minimum of 30 ***unique*** credit hours in their major and 20 ***unique*** credit hours in their minor for a total of at least 50 unique credit hours between them. Some majors that require 50 or more hours themselves do not require a minor; students should check requirements of the selected major in the undergraduate catalog to see if a minor is required.
- Earn no more than 60 credit hours in one subject area (prefix). Credits in excess of the 60 maximum will not be counted toward the minimum of 124 credits required for a bachelor's degree.
- Earn the minimum number of credits in 300-level and above courses in each major and minor as specified below - these credits must be earned in distinct courses; that is, no course can be used to fulfill this requirement in more than one major or minor.
    - Earn a minimum of 6 credits in 300-level or higher courses at EMU in each minor
    - Earn a minimum of 9 credits in 300-level or higher courses at EMU in each major that requires a minor.
    - Earn a minimum of 15 credits in 300-level or higher courses at EMU in each major that does not require a minor
- Transfer credit will be awarded for courses taken at colleges and universities that are accredited by one of the recognized regional accrediting bodies only if the courses are college-level (equated to 100-level or above at EMU) and the student earned a "C" (or 2.0 on a 4 point scale) or better. Transfer credit may be awarded on a case-by-case basis for college-level courses in which a "C" (2.0) or better was earned at institutions outside the U.S. or at non-accredited U.S. institutions; the internal review of such courses is conducted by individual departments/schools within EMU, and additional documentation may be required. *Please note:* EMU awards only credits for transferred courses; grades are not used in the calculation of an EMU GPA.
- Earn a minimum cumulative GPA of 2.0 in courses taken at EMU in order to graduate. In addition, a minimum cumulative GPA of 2.0 must be reached in each major and minor. Only courses taken at EMU and those applied to a student's major or minor will be used in the calculation of their major and minor cumulative GPAs. (Note: some

programs may require a higher GPA - check with your program advisor.)

General Education Requirements EMU's General Education Program requires students to choose from a menu of approved courses in several different areas; do not assume that other courses in the same department or with similar names will fulfill these requirements. A detailed description of General Education requirements is available in the General Education section of the catalog.

Students who transferred to EMU may have modified general education requirements based on Michigan Transfer Agreement (MTA) or articulation agreements; consult your academic advisor for additional information.

EASTERN MICHIGAN UNIVERSITY
DIVISION OF ACADEMIC AFFAIRS
COLLEGE OF ARTS AND SCIENCES

OUTLINE FOR SUBMITTING PROPOSALS FOR NEW INTERDISCIPLINARY DEGREE PROGRAMS

Use this outline to prepare proposals for new interdisciplinary programs, including undergraduate majors and minors and graduate majors. Proposals should be submitted in narrative form, organized according to the following outline. Guidelines for submitting such proposals are on the following pages.

PROPOSED INTERDISCIPLINARY PROGRAM NAME: ____DATA SCIENCE AND ANALYTICS (DSA)_____

DEGREE:__Bachelor____ REQUESTED START DATE: ____Fall 2016_____

PROGRAMS/DEPARTMENT(S)/SCHOOL(S):_Computer Science__

PROGRAMS/DEPARTMENT(S)/SCHOOL(S):_Mathematics__

PROGRAMS/DEPARTMENT(S)/SCHOOL(S):_Computer Information Systems__

PROGRAMS/DEPARTMENT(S)/SCHOOL(S): _____

COLLEGE(S): _ARTS & SCIENCES__

COLLEGE(S):_BUSINESS_

CONTACT PERSON:__Andrew Ross, Mathematics Dept._____ CONTACT PHONE: 734-487-1658

CONTACT EMAIL:ANDREW.ROSS@EMICH.EDU

ALSO: **STEVAN MRDALJ (CIS) 734-487-2671** SMRDALJ@EMICH.EDU ,

WILLIAM SVERDLIK (COMP.SCI.) 734-487-7081 WSVERDLIK@EMICH.EDU

I. Description:

A. Goals, Objectives, Student Learning Outcomes

The Data Science and Analytics (DSA) program is an interdisciplinary program designed to provide broad knowledge and the technical competency required to manage and analyze the ever-growing body of data in industry, government, science, and nonprofit organizations.

The program is interdisciplinary and the development of the program involved several departments in two different colleges. Other departments are allowed to request affiliation with the DSA program, since our market research indicates that this field is truly interdisciplinary. This program will be governed according to the Program Input Document supplied in Appendix A.

Graduates of our program will be able to organize, aggregate and analyze rapidly generated data from diverse sources. Such skills are already in high demand in such diverse disciplines as econometrics, bioinformatics, remote sensor networks, and health care analytics. Demand for highly qualified data scientists in these fields will only grow in the future, providing opportunities for our graduates to contribute to scientific research, and engage within the community at the local and global level.

Our program is both practical and theoretical, providing students the skills to contribute immediately upon graduation and the ability to assimilate emerging technologies as they develop. Graduates of our program must be life-long learners; this is consistent with the goals of our current constituent departments within the program. This is consistent with the stated vision and mission statement of the College of Arts and Sciences.

1.1.1. Promote and create Interdisciplinary, interactive learning events and courses 1.1.2. Create incentives for faculty to engage in interdisciplinary work"

The program is aligned with the EMU Strategic Plan; for example, the program includes a capstone project course that is part of "*Objective 1.1:* Document, review and continue to expand field experiences (internships, practicum, co-op activities),"

All DSA students need to complete the common core requirements which are designed to provide the foundational theory and skills required for Data Science and Analytics. Students can tailor their curriculum by following one of three concentrations to gain a deeper understanding of a data science issue depending on their interest. Currently, concentrations in Data Science Theory, Web Analytics and Advanced Data Analysis are available, but more can be developed in the future. The program is consistent with university strategic direction in that it provides practical experience in the form of a capstone experience that enhances learning and helps address community needs through internships or research.

The main source of employment for graduates of the DSA program would be companies in any sector of industry, research institutions, regulatory agencies, nonprofit organizations, and consulting firms. Although many of the jobs available to students who complete the DSA program are entry-level, experience and graduate work can allow students to advance to management positions.

Upon graduating from the DSA program, students will have achieved the following learning outcomes:
SLO A: Apply business concepts and methods to define and structure problems.
SLO B: Apply statistics concepts and methods to design studies and analyze data.
SLO C: Apply computer science concepts and methods to organize data.
SLO D: Disseminate their work to technical and nontechnical people.
SLO E: Learn new needed skills on their own as lifelong learners.

B. Program

Data Science and Analytics Program Structure

General Education Requirements: approx. 41 hours

For specific requirements refer to the General Education Program .

Major Requirements: 67 or 68 hours

Required Courses: 55 hours

- MATH 120 - Calculus I (GEQR) 4 hrs
- MATH 121 - Calculus II 4 hrs
- MATH 122 - Elementary Linear Algebra 3 hrs

- COSC 146 - Applied Programming 3 hrs
- COSC 111 - Introduction to Programming 3 hrs
- COSC 112 - Introduction to Programming Online Lab 1 hr
- COSC 211 - Programming Data Structures 3 hrs
- COSC 212 - Programming Data Structures Online Lab 1 hr
- COSC 311 - Algorithms and Data Structures 3 hrs
- COSC 472 - Introduction to Big Data 3 hrs
- COSC 481W - Software Engineering and Senior Project (GEWI) 3 hrs

- IS 385 Data Warehousing, 3 hrs (new course)
- IS 410 – Data Mining 3 hrs (title change)

- MATH 360 - Statistical Methods 3 hrs

One course from the following:

- MATH 461 - Linear Regression Analysis 3 hrs
- DS 465 - Applied Linear Statistical Models 3 hrs

One course from the following:

- COSC 231 - Internet-based Computing 3 hrs
- IS 247 - Introduction to Web Application Development 3 hrs (*)

One course from the following:

- IA 212 - Open Systems Platform and Network Administration
- IS 416 - System Software Administration (*)

One course from the following:

- COB 200L4 - Introduction to Business (GELB) 3 hrs
- IS 350 - Enterprise Resource Planning and Architecture 3 hrs (*)

One course from the following:

- IS 380 – Data and Information Management 3 hrs (*) (New title effective Winter 2106)
- COSC 471 - Database Principles 3 hrs

* these IS courses would have the IS 215 prerequisite waived for DSA majors

Concentration Courses: 12 or 13 hours

Select one of the following concentrations

Concentration A (Theory) Courses (12-13 hours)

Choose four courses from the following:

- MATH 223 - Multivariable Calculus 4 hrs
- STAT 370 - Probability and Statistics I 3 hrs
- MATH 419W - Introduction to Stochastic Mathematical Modeling (GEWI) 3 hrs
- COSC 221 - Computer Organization I 3 hrs
- COSC 423 - Computer Operating Systems 3 hrs
- COSC 439 - Computing Network Principles 3 hrs

Concentration B (Web Analytics) Courses (12 hours)

- IS 310 - Social Media for Business 3 hrs
- IS 339 - Google Analytics and Landing Page Optimization 3 hrs
- IS 405 Web Content Mining, 3 hrs (new course)
- IS 422 Social Media Network Analysis, 3 hrs (new course)

Concentration C (Advanced Data Analysis) Courses (12 hours)

Choose four courses from the following:

- COSC 462 Introduction to Information Retrieval, 3 hrs
- COSC 473 Big Data II, 3 hrs
- STAT 462 - Design and Analysis of Experiments 3 hrs
- STAT 474W - Applied Statistics (GEWI) 3 hrs
- STAT 468 - Introduction to Biostatistics 3 hrs
- MATH 419W - Introduction to Stochastic Mathematical Modeling (GEWI) 3 hrs

Minor Requirements:

This program does not requires a minor.

Program Total:

Students must earn a minimum total of 124 credits at the 100-level or above; 67 to 68 of those are due to this program.

All new course proposals appear in a separate electronic folder along with this proposal.

A sample typical program of study that a student could follow in completing the DSA program is given in Table 1. For the pairs of courses where the student chooses one of the pair, we only list the first one of the pair (which usually has a lower course number, e.g. DS 265 versus Math 360). The "Advanced Data Analysis" concentration is used as an example; the student only needs to choose 4 of the 6 courses in the concentration, but this example shows that it is even possible to take all 6 of them if the student is inclined.

Prefix'	Number'	Title'	Credits'	Semester'
COB'	200L'	Introduction'to'Business'	3'	1'
COSC'	146'	Applied'Programming'	3'	1'
MATH'	120'	Calculus'I'(also'satisfies'GenEd'QR)'	4'	1'
COSC'	111/112'	Intro'to'Programming'	4'	2'
MATH'	121'	Calculus'II'	4'	2'
MATH'	122'	Elementary'Linear'Algebra'	3'	2'
COSC'	231'	InternetDbased'Computing'	3'	3'
COSC'	211/212'	Programming'Data'Structures'	4'	3'
IA'	212'	Open'Systems'Platform'and'Network'Administration'	3'	3'
COSC'	311'	Algorithms'and'Data'Structures'	3'	4'
DS'	265'	Business'Statistics'I'	3'	4'
COSC'	472'	Big'Data'	3'	5'
IS'	380'	Data'and'Information'Management*'	3'	5'
MATH'	461'	Linear'Regression'Analysis'	3'	5'
IS'	385'	Data'Warehousing"'	3'	6'
IS'	410'	Data'Mining*'	3'	6'
MATH'	419W'	Introduction'to'Stochastic'Mathematical'Modelling'	3'	6'
COSC'	479'	Information'Retrieval'and'Recommender'Systems'	3'	7'
COSC'	4xx'	Big'Data'II'	3'	7'
MATH'	462'	Design'and'Analysis'of'Experiments'	3'	7'
COSC'	481W'	Software'Engineering'and'Senior'Project'	3'	8'
MATH'	468'	Introduction'to'Biostatistics'	3'	8'
MATH'	474W'	Applied'Statistics'	3'	8'

* new title effective Winter 2016

C. Admission

No special admission requirements beyond those required for admission to the University

D. Projections

Given the justification presented in section II it is reasonable to expect an annual enrollment of 20 to 30 students once it is fully developed. If a more coordinated marketing effort is developed and implemented, the initial

II. Justification/Rationale

Analyzing large data sets—so called big data—has become a key basis of both science and business, underpinning new waves of productivity growth, innovation, and consumer behavior analytics. To remain competitive and maximize efficiencies in the coming decade, companies need to invest in big data technologies. Equally important will be the investment in data scientists' proficiency in advanced statistics and machine learning. Nobody doubts that companies will continue to ramp up hiring, recruitment and training of data scientists. But there seems to be a growing alarm that we won't have enough data scientists to go around.
Tom Davenport and D.J. Patil's Harvard Business Review article, "Data Scientist: The Sexiest Job of the 21st Century," (https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century/ar/1) offers a detailed look at the role of the data scientist and the requisite skills for the position. They define data scientists as skilled workers who combine data, analytic, statistic and computer science capabilities, and who can clearly define requirements and communicate effectively with line-of-business executives and corporate management. While this new breed is a critical ingredient in the wave of Big Data initiatives, they declared that finding them is incredibly challenging and the shortage of data scientist is becoming a serious constraint in many industry sectors. Tom Davenport, author of the highly acclaimed best seller Competing on Analytics: The New Science of Winning, asserts "There is a storm approaching on the Big Data talent front and there is no reliable source of talent in this emerging, fast growing category. It would seem to be a wise move to begin building such talent as well as buying it".

Paul Barth and Randy Bean in their Harvard Business Review article "There's No Panacea for the Big Data Talent Gap" (https://hbr.org/2012/11/the-big-data-talent-gap-no-pan) surveyed senior Fortune 500 and federal agency business and technology leaders to discover the level of serious interest surrounding Big Data. The survey of C-suite and executive function heads with responsibility for Big Data initiatives revealed that 85% of the organizations surveyed had funded Big Data initiatives or were in the planning stage. The interest and commitment is real. What is less certain is how these same organizations plan to support these initiatives from a business and talent perspective. Their survey indicates there will be plenty of competition for their talents. The following are responses on the two questions on how challenging it is to source analytical skills.

How challenging is it to source analytical skills in general?

How challenging is it to source data scientists?

Similarly, the NewVantage Partners (the premier consultancy focused on guiding business and technology executives in the creation of strategic solutions that harness the power of their data) surveyed (http://newvantage.com/wp-content/uploads/2012/11/NVP-Press-Release-Big-Data-Talent-Survey-111312.pdf) perspectives from Fortune 1000 C-Suite executives and Federal government leaders on:

- The current state of Big Data and analytics talent and capabilities across their organizations
- The talent gap that they are trying to fill – a critical combination of scientific background, computational and analytics skills
- How their organizations are recruiting and developing talent for Big Data initiatives
- The growing need for IT and business leaders who understand how to optimize Big Data Opportunities

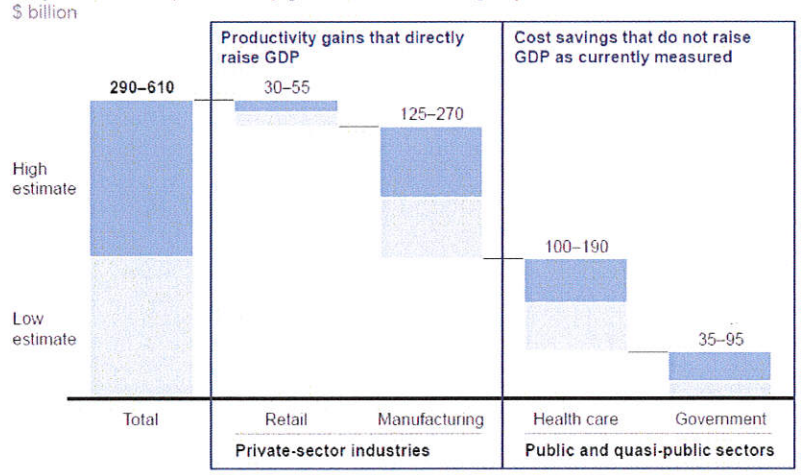
Randy Bean, co-founder and managing partner of NewVantage Partners, states in this report "Simply put, Big Data talent is

- 37% of respondents ranked their current analytics capabilities as less than adequate
- 70% of respondents are hiring or plan to hire Data Scientists in the near future – with 60% stating that it is very difficult to find and hire these individuals
- 91% of respondents are hiring new people from outside their organization to fill the talent gap while 69% are training their existing analytics professionals
- 51% stated that it is challenging to find IT professionals with strong data management skills and familiarity with new technologies that apply to Big Data
- 50% reported that it is very difficult to find and hire business leaders and managers who can identify and optimize business opportunities in Big Data.
- Only 2% reported that they have had no challenges

The McKinsey Global Institute (MGI) estimates 40,000 exabytes of data being collected by 2020 — up from 2700 exabytes in 2012 As a result, the opportunity for data scientists and organizations that move early in the space is apparent. MGI also published a report (http://www.mckinsey.com/features/big_data) that by 2018 the United States will experience a shortage of 190,000 skilled data scientists, and 1.5 million managers and analysts capable of reaping actionable insights from the big data deluge. This research also examines the state of digital data and documents the significant value that can potentially be unlocked. This trend receives additional support in subsequent report "Game changers: Five opportunities for US growth and renewal", (http://www.mckinsey.com/insights/americas/us_game_changers) by predicting incredible potential GDP growth driven by Big Data Analytics displayed in the following figure.

Big data analytics can raise US GDP by up to $325 billion in retail and manufacturing, and create up to $285 billion in cost savings in health care and government

Projected annual productivity gains and cost savings by 2020
$ billion



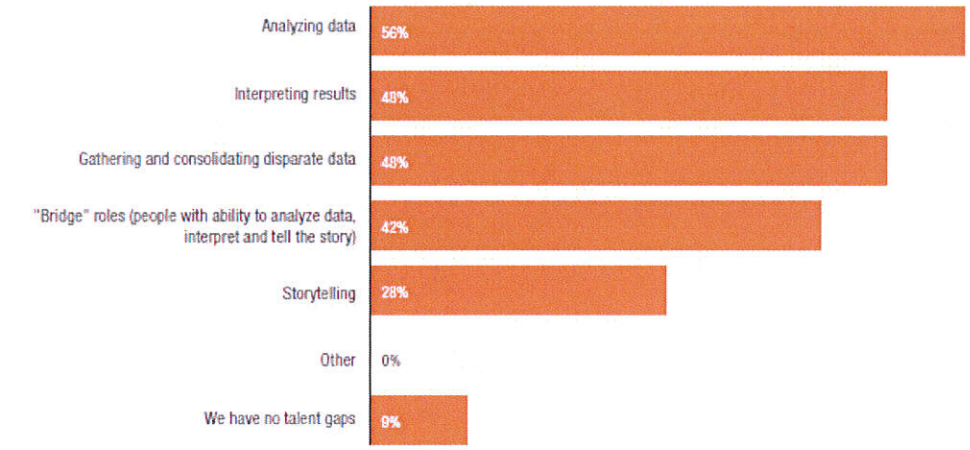
SOURCE: McKinsey Global Institute analysis

MGI cites the Obama administration's Big Data Research and Development Initiative as one promising response to this economic challenge, and predicts that "The adoption of big data tools in the federal government may provide additional momentum in academia, the biomedical research community, and the private sector."

The value of data science skills, and resulting demand for skilled practitioners, is a well-known challenge within the emerging big data analytics industry. The online employment and career site Dice reports on the demand for big data skills is expanding across the United States and in diverse industries. It calls this trend that refers to as "big data sprawl." According to this report (http://www.eweek.com/it-management/slideshows/10-tips-to-taking-the-stingand-penaltiesout-of-a-software-license-audit.html), the number of big data jobs on the site has shown a triple-digit surge year-over-year. The largest percentages of big data jobs on Dice are in Silicon Valley, New York and Washington, D.C./Baltimore. "Big data" has become a household word, and many well-known companies are seeking big data talent, the Dice report said. The types of companies looking for IT pros with big data skills include consulting firms, defense contractors, e-learning businesses, financial firms, online dating sites and more.

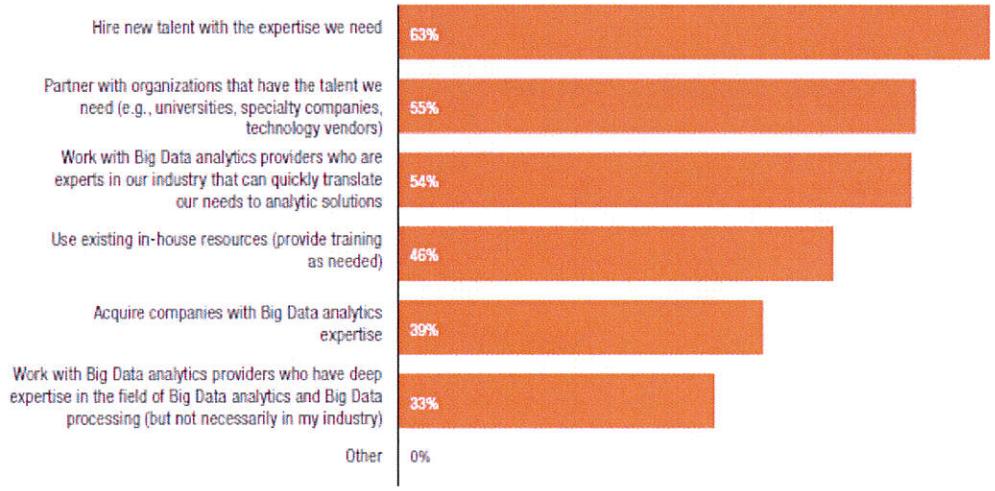
Executives surveyed by Accenture (one of the world's leading professional services companies) are aware of their own big data talent shortage and understand the critical nature of sourcing and developing the talent needed to close the big data talent gap (https://www.accenture.com/us-en/insight-industrial-internet-competitive-landscape-industries.aspx). About half of those surveyed note that they have talent gaps in several critical areas including analyzing data, interpreting results and gathering and consolidating disparate data as indicated in the following figure.

In which of the following areas do you have gaps in your talent? (Multiple responses)



Hiring talent with the expertise needed is the most obvious remedy to the talent gap issue, named by 63 percent of survey respondents. Yet the fact is that there won't be enough experienced talent. Indeed, shortages in the number of data scientists are projected, as well as the number of managers capable of using Big Data analyses to make good decisions.

Which of the following will your company pursue to ensure you have the talent needed to fulfill your Big Data analytics strategy?



The results of two surveys about data scientists were released this week, covering both the supply and demand sides of this emerging new profession:

The MIT Sloan Management Review and SAS 2015 annual survey (http://sloanreview.mit.edu/projects/analytics-talent-dividend) reports on the challenges of recruiting, training, and integrating data scientists. This survey of 2,719 business executives, managers and analytics professionals worldwide presents the specific challenges of acquiring and benefiting from this still-scarce talent pool of "the sexiest job of the 21st Century." Four in ten (43%) companies report their lack of appropriate analytical skills as a key challenge but only one in five organizations has changed its approach to attracting and retaining analytics talent. As a result of the scarcity of data scientists, 63 percent of the companies surveyed are providing formal or on-the-job training in-house. These companies are also doing more to train existing managers to become more analytical (49%) and train their new data scientists to better understand their business (34%). Still, half of the survey respondents cited turning analytical insights into business actions as one of their top analytics challenges.

1. **Data crunching**. The era of big data is just getting started, with many firms eager to tap vast new databases to gather more info on their customers, their competitors, and even themselves. The challenge isn't just crunching numbers; it's making sense of them, and gaining useful insights that can be translated into a business edge. Marketing and market research are two growing fields where the use of data is exploding.

Finding the right Big Data talent and skilled resources finding and retaining Big Data talent will become even tougher as competition for this limited talent pool intensifies over time. Perhaps the best indicator of the need for these skills is the number of data science programs springing up around the country.

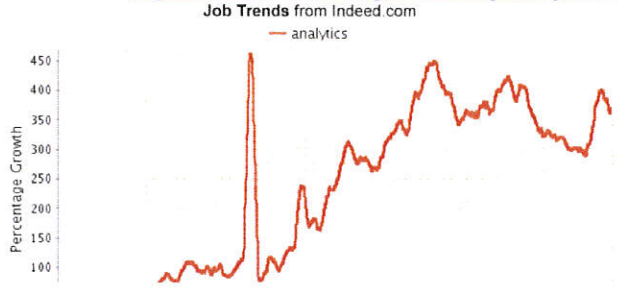
Job Trends

On August 5th, 2015, we queried the job posting site indeed.com to show the trends in four key words: "data science", "big data", Hadoop (a key big-data software package), and "analytics"; the graphs show strong trends:
http://www.indeed.com/jobtrends?q=%22data+science%22&l=&relative=1

http://www.indeed.com/jobtrends?q=hadoop&l=&relative=1

http://www.indeed.com/jobtrends?q=%22big+data%22&l=&relative=1

2015-08-05 http://www.indeed.com/jobtrends?q=analytics&l=&relative=1



Survey of Undergraduate Degrees in Data Science

It is difficult to exhaustively determine comparable degrees in Data Science, as the term "Data Science" is not standard. Alternate degree names include "Data Mining" (Central Michigan University), "Analytics"(North Carolina State University), "Business Analytics", (Arizona State University, Bentley University, Drexel University), Data Analytics Engineering (George Mason University), Predictive Analytics" (DePaul University). Both Western Michigan University and the University of Michigan have introduced programs entitled "Data Science". Often, but not always, the inclusion of the term "analytics" in the program name implies participation of a business college. Many of the aforementioned programs are for certificates or masters degrees.

We note that Washtenaw Community College now offers a certificate in "Applied Data Science" (see http://www.wccnet.edu/academics/programs/view/program/CTADS/).

A careful review of many of the existing programs reveals a common practice: create new degrees by combining existing courses from several different departments. Frequently, no new courses (and hence, no new technologies) are introduced. EMU distinguishes itself from these programs with the inclusion of several new courses that include the latest in parallel/distributed technologies.

Since this proposal is for an undergraduate degree, we attempt to survey existing undergraduate programs. The web page

http://101.datascience.community/2013/08/21/undergraduate-programs-in-data-science/

lists the following undergraduate degrees in Data Science as of August 21, 2013:

- College of Charleston – B.S. in Data Science
- University of San Francisco – B.S. in Data Science
- University of Iowa – B.A. in Business Analytics
- Northern Kentucky University – B.S. in Data Science
- University of Warwick – BSC Data Science
- Illinois Institute of Technology – B.S. in Computer Science with specialization in Data Science
- George Mason University – Data Science major & minor
- Northwestern College in Iowa – Online Analytics certificate
- The Ohio State University – B.S. Data Analytics

The magazine of the American Statistical Association provides links to newer programs:

http://magazine.amstat.org/blog/2015/07/01/new-undergraduate-data-science-programs/

http://magazine.amstat.org/blog/2015/08/01/new-undergraduate-data-science-programs-2/

The American universities include

University of California, Irvine – B.S. in Data Science
Winona State University – B.S in Data Science
University of Michigan – B.S. in Data Science
Miami University (Ohio) – Analytics

We are unable to determine if the proposal from Western Michigan University has been approved.

We look at some current undergraduate degrees:

1)University of San Francisco

Welcome

Get involved in the emerging field of "big data" with the Bachelor of Science degree in data science (BSDS) at the University of San Francisco. This challenging and interdisciplinary major provides rigorous academic training in mathematics and quantitative skills, programming, and problem solving for data-intensive fields such as economics, biology, computer science, and many others.
The core courses in the BSDS major are in mathematics and computer science, with 52 units distributed among these two

- mathematical data science (12 units)
- computational data science (12 units)
- economic data science (12 units)

for a total of 68 units.
This is a large major, but comparable with other science majors such as Chemistry and Biology. As with those majors, this is intended for focused, well-prepared students who are interested in spending the bulk of their college career developing skills within their major.

2) College of Charleston (Core data science curriculum)

Data Science Major Requirements

Degree: Bachelor of Science
Credit Hours: 69+
"PR" indicates a pre-requisite. "CO" indicates a co-requisite.
Courses within this major may also satisfy general education requirements. Please consult http://registrar.cofc.edu/general-edu for more information.
Required Courses

DATA 101	Introduction to Data Science (3) PR: None
DATA 210	Dataset Organization and Management (3) PR: None
DATA 495	Data Science Capstone (3) PR: DATA 210, CSCI 470, and MATH 441

Math Requirement

MATH 120	Introductory Calculus (4) PR: C- or better in MATH 111 or placement
MATH 203	Linear Algebra (3) PR: MATH 220 or instructor permission
MATH 207	Discrete Structures I (3) PR: MATH 105, MATH 111, or MATH 120
MATH 220	Calculus II (4) PR: MATH 120 or HONS 115
MATH 250	Statistical Methods (3) PR: MATH 111 or MATH 120 or instructor permission
MATH 350	Statistical Methods II (3) PR: MATH 120, MATH 250
MATH 440	Statistical Learning I (3) PR: MATH 203, MATH 220, MATH 350
MATH 441	Statistical Learning II (3) PR: MATH 440

Computer Science Requirement

CSCI 220	Computer Programing I (3) PR: CSCI 120 or CSCI 180 or CSCI 210 or MATH 111 CO: CSCI 220L
CSCI 220L	Computer Programing I Lab (1) CO: CSCI 220
CSCI 221	Computer Programing II (3) PR: CSCI 220, CSCI 220L; CO or PR: MATH 207
CSCI 230	Data Structure and Algorithms (3) PR: CSCI 221, MATH 207
CSCI 310	Advanced Algorithms (3) PR: CSCI 230, MATH 207
CSCI 334	Data Mining (3) PR: CSCI 221, MATH 207, MATH 250
CSCI 470	Principles of Artificial Intelligence (3) PR: CSCI 230, MATH 307

3)Illinois Institute of Technology (BS in Computer Science with Specialization on Data Science)

Specialization in Data Science

Four courses are required for the Specialization in Data Science:

- CS 422 Data Mining or CS 584 Machine Learning
- CS 451 Intro to Distributed Computing (CS 495-05 in Fall 2012)
- BUS 371 Strategies for Reaching New Markets
- MATH 481 Introduction to Stochastic Processes or MATH 483 Design and Analysis of Experiments
    - MATH 481 Prerequisites: MATH 332/333 and MATH 475.
    - MATH 483 Prerequisite: MATH 476.

III. Preparedness

A. Qualifications of Faculty

Most of the courses in the DSA program are pre-existing and will be taught by faculty vetted by the departments that controls those courses. The new proposed courses are also developed by existing faculty from various departments who are qualified by virtue of their experience in teaching and/or research in the topic area. The initial faculty affiliates involved in developing this proposal are listed below:

Mathematics/Statistics Faculty:

- Andrew Ross, Professor, Ph.D., Industrial Engineering and Operations Research
- Tanweer Shapla, Associate Professor, Ph.D., Statistics
- Rita Chattopadhyay, Professor, Ph.D. Statistics
- Kathy Chu, Professor, M.P.H, Ph.D. Statistics
- John Curran, Professor, Ph.D. Mathematics

Computer Science Faculty:

- William Sverdlik, Professor, Ph.D., Computer Science
- Susan Haynes, Professor, Ph.D., Computer Science
- Krish Narayanan, Professor, Ph.D., Computer Science
- Li Zhang, Associate Professor, Ph.D., Computer Science

Computer Information System Faculty:

- Stevan Mrdalj, Professor, Ph.D., Computer Information Systems
- Yaman Roumani, Assistant Professor, Ph.D., Computer Information Systems
- David Chou, Professor, Ph.D., Computer Information Systems

Decision Sciences Faculty:

- Joseph Scazzero, Professor, Ph.D., Decision Sciences

B. Library Resources

We performed a search of the online database of Halle Library for "Hadoop, MongoDB, CouchDB and HBase" ; these are some of the leading technologies currently in the Data Science community. In all cases, the library has access to multiple sources on each topic including online ebooks. Assuming that these online references remain, we do not anticipate extraordinary library resources. Current access to sources such as Books 24x7 and the O'Reilly book and media series are sufficient.

C. Adequacy of Existing Facilities/Equipment

Since all of courses in the program are housed in individual departments, it is largely the responsibility of the departments to assure adequacy of facilities/equipment for the courses. In addition, the proposed program will utilize the following existing labs and equipment:

- Computer Information Systems computer classrooms
- Computer Information Systems lab
- Computer Science computer classrooms
- Statistical computer classrooms

D. Not applicable to this proposal

E. Marketing Plan for the New Program

A website that provides information and explains the requirements of the new major will be developed. The

are candidate feeder schools for the departments that are participating in the new major. Pamphlets will be prepared and sent to these schools. The program coordinator will visit a number of these schools on special 'college nights' to discuss the new program with guidance counselors, teachers, and interested students. A demonstration booth will be prepared to market the program to current and prospective EMU students at university-sponsored programs such as 'Explore Eastern'. This marketing campaign will begin as soon as the program is approved.

IV. Assessment/Evaluation

A) Student Learning Outcomes

Because data science is a field that changes rapidly with advances in technology, we have been careful to phrase our list of SLOs in generic language rather than specifying particular software packages, for example. For the same reason, we are including as a goal the student's ability to teach themselves new things (like new software) as trends change after graduation (or perhaps even during their undergraduate program).

Upon graduating from the DSA program, students will have achieved the following learning outcomes:

SLO A: Apply business concepts and methods to define and structure problems.
SLO B: Apply statistics concepts and methods to design studies and analyze data.
SLO C: Apply computer science concepts and methods to organize data.
SLO D: Disseminate their work to technical and nontechnical people.
SLO E: Learn new needed skills on their own as lifelong learners.

B) Curriculum Mapping

Here is how each course relates to the Student Learning Outcomes. The table is organized by knowledge domains first and then rough chronological order, rather than by chronological order first.
In this table, a 1 signifies beginner, 2 signifies intermediate, and 3 signifies advanced.

''	A:' Business'	B:' Statistics'	C:' CompSci'	D:' Communication'	E:'Lifelong' Learning'
MATH'120'		1'			
MATH'121'		1'			
MATH'122'		2'			
DS'265'		2'		1'	
MATH'461'		3'		2'	
COSC'146'			1'		1'
COSC'111/112'			1'		1'
COSC'211/212'			2'		2'
COSC'311'			2'		
COSC'231'			2'		2'
COSC'472'			3'	3'	3'
COSC'481W'			3'	3'	3'
''					

"					
COB'200L'	1'			1'	
IS380'	2'		2'		1'
IS385'	2'				2'
IS410'	3'			3'	3'

C) Monitoring Program Objectives through Assessment Data

A subcommittee of the steering committee will organize the assessment efforts.

Instructors of the classes listed as a "3" in the above table will be supplied with a list of the DSA majors in their class, and they will return to the DSA program graded copies of SLO-relevant assessments (tests, reports, presentations) for those students. The capstone project course will be the main source of information on whether students are meeting the SLOs by the time of graduation, but we will collect data on courses before that to help catch potential problems before it would be too late to fix them.

By definition, it is difficult to assess SLO E (Lifelong Learning) during an individual's undergraduate curriculum. We have some experience already in the Big Data course, asking students to teach themselves a programming language that was not a prerequisite with only minimal in-class support; their results were then assessed in a homework assignment. We will also gather information about it by asking each DSA major about it with a survey to be developed, about how confident they are that they can teach themselves new skills. While this has the potential to give overestimated abilities, it could give us a warning if the students themselves say they are not very confident in their abilities. We will also reach out to alumni a year after graduation and survey them about their impressions of the program, aligned with the SLOs.

D) Timeline for Monitoring Program (projected)

Since most of the courses in the DSA program already exist, assessment can begin immediately upon the entry of DSA majors to the courses tagged for assessment above. Each year, the assessment report will be given to the Faculty Affiliates for consideration and discussion of improvement strategies.

E) Developing Program-level Outcomes

The program coordinator will regularly meet with a subcommittee of faculty affiliates and relevant outside parties (local employers, etc.) to determine if the program-level outcomes and their assessments need to be updated based on industry trends. They will also discuss high-level indicators such as student progression through the program, % of program graduates going to various fields (industry, nonprofit, government, graduate school, unemployed, etc.), and so forth.

A. Faculty, lecturers or supportive staff required.

The program will require a Program Coordinator with a ½ time course release for Fall and Winter and 20% of base salary for Summer semesters. It is essential to have a Program Coordinator available for Summer due to the number of students who seek information and advice during that time span. The actual cost will be determined when this person is selected.

B. Space or facilities required.

No additional space for the program is needed.

C. Equipment required.

Many of the "Big Data" vendors provide free distributions of software to run on virtual machines (e.g. Cloudera). In addition, software can now run "in the cloud" on user configurable distributed systems (e.g. Amazon Web Services). This means that the University will incur no extra costs associated with software or hardware other than the present costs associated with University computer labs. We note that classes currently employing the cloud model provide ample exposure to highly sophisticated modes of computing at a cost of under $50 per student per class. This cost will be paid directly by the student.

D. Assistantships/fellowships required.

Initially the program will need a part-time graduate student assistant to help with the work of the Program Coordinator.

E. Library resources required.

Since this is a fast-moving field, we need to maintain access to sources such as Books 24x7 and the O'Reilly book and media series.

F. Marketing and recruiting costs.

Suggested budget of $5000/year for the first three years to cover website development, High School recruitment, University Fairs, brochures, travel expenses and other related promotional costs.

G. Other costs not covered above.

This is a rapidly changing field; faculty require professional development. We would anticipate 6 faculty (2 from each of mathematics, computer science, and information systems) attending conference/tutorials/workshops per year. Conference registration fees tend to be higher in this field than math conferences, for example.

H. Total of all financial requirements for implementation of proposed degree.

Total costs cannot be exactly determined until the Program Coordinator is selected, courses are offered, enrollment numbers are available, and recruitment plans are established. The following is a draft summary of estimated expenses per year for the first 3 years (see descriptions above for justification):

	Each'	Qty'	Line'total'
1'course'release,'cost'to'replace'with'an'adjunct'	$3,600"	4'	$14,400"
Summer'salary,'20%'of'approx.'$70,000'	$14,000"	1'	$14,000"
Marketing'(only'first'3'years)'	$5,000"	1'	$5,000"
Professional'development'conferences'	$1,200"	6'	$7,200"
			$40,600"

I. Continuing Education/Extended Programs:

VI. Action of the Program(s)/Department(s)/College(s)

(Include the faculty votes signatures from all collaborating programs/departments/schools, if applicable.)

1. Program/Department/School (Include the faculty votes signatures from all submitting programs/departments/schools.)

a. Vote of program faculty:_MATHEMATICS ____ For_20_ Against_0 ____ Abstentions_0 _____

I support this proposal. The proposed program can __X__ cannot _____ be implemented without additional College or University resources.

[redacted]

Program Administrator Signature — Date: 9/28/15

b. Vote of department/school faculty:_COMPUTER SCIENCE__For_11 ____ Against 0____Abstentions _1

I support this proposal. The proposed program can __X__ cannot _____ be implemented without additional College or University resources.

[redacted]

Department Head/School Director Signature — Date: 9/28/15

c. Vote of department/school faculty: COMPUTER INFORMATION SYSTEMS For_12_Against_0___Abstentions_0_

I support this proposal. The proposed program can __Y__ cannot _____ be implemented without additional College or University resources.

[redacted]

Department Head/School Director Signature — Date: 10/26/2015

2. College(s)/Graduate School (Include signatures from the deans of all submitting colleges.)

A. College: [redacted]

I support this proposal. The proposed program can __N__ cannot _____ be implemented within the affected College without additional University resources.

College Dean Signature — Date: 11/10/15

B. College: ARTS AND SCIENCES

I support this proposal. The proposed program can __X__ cannot _____ be implemented within the affected College without additional University resources.

[redacted]

College Dean Signature — Date: 11/13/2015

VII. Approval

Associate Vice-President for Academic Programming Signature — Date

VIII. Appendices

Data Science and Analytics (DSA)

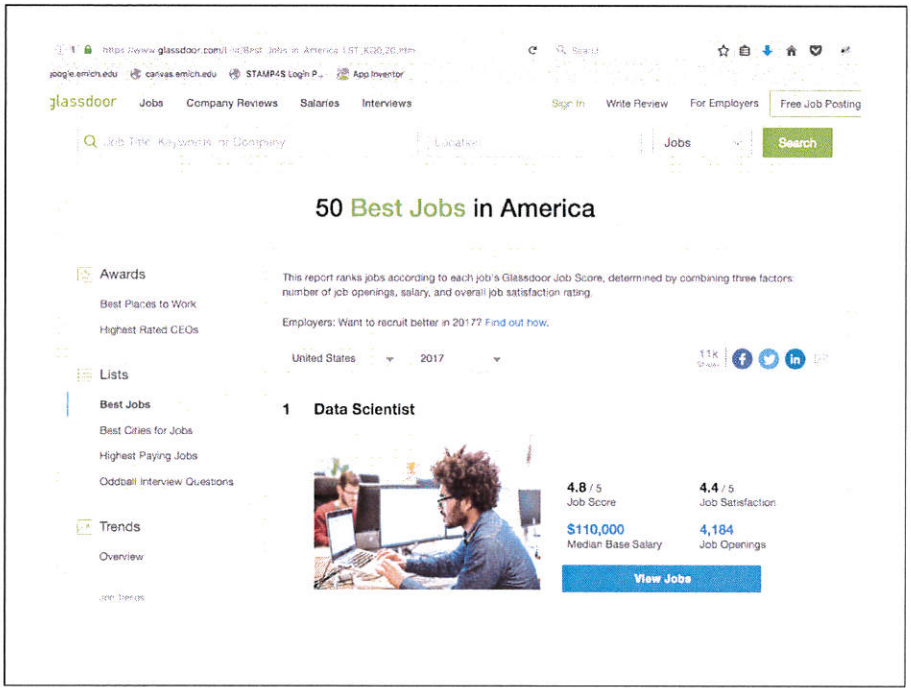
A New Interdisciplinary Program

Presented to the

Board of Regents

Eastern Michigan University

April 21, 2017



Data Science and Analytics (DSA)

Help Wanted: Black Belts in Data

Starting salaries for data scientists have gone north of $200,000

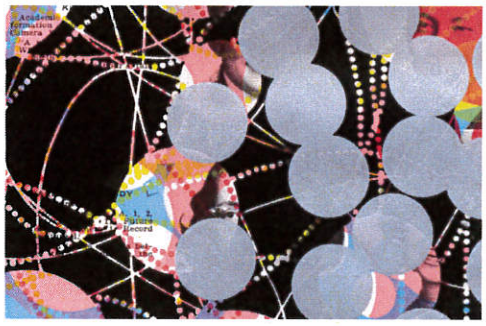
by **Rodrigo Orihuela** and **Dina Bass**
June 4, 2015, 1:07 PM EDT *Updated on* June 4, 2015, 2:00 PM EDT
From **BloombergBusinessweek** | Subscribe | Reprints

KDnuggets Home » News » 2016 » Jan » News, Features » Busin

Businesses Will Need One Million Data Scientists by 2018

Data Science and Analytics - DSA

Harvard Business Review



DATA

Data Scientist: The Sexiest Job of the 21st Century

by Thomas H. Davenport and D.J. Patil

What Do Data Scientist Do?

Data Science Program Get Started Current St

In simple terms, a data scientist's job is to analyze data for actionable insights.

Specific tasks include:

- Identifying the data-analytics problems that offer the greatest opportunities to the organization
- Determining the correct data sets and variables
- Collecting large sets of structured and unstructured data from disparate sources
- Cleaning and validating the data to ensure accuracy, completeness, and uniformity
- Devising and applying models and algorithms to mine the stores of big data
- Analyzing the data to identify patterns and trends
- Interpreting the data to discover solutions and opportunities
- Communicating findings to stakeholders using visualization and other means

Data Science and Analytics (DSA)

So what ?

We've been doing this for a long time!

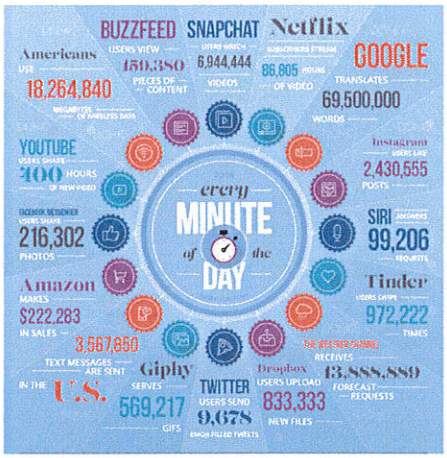
What's different ?

Data Science and Analytics (DSA)

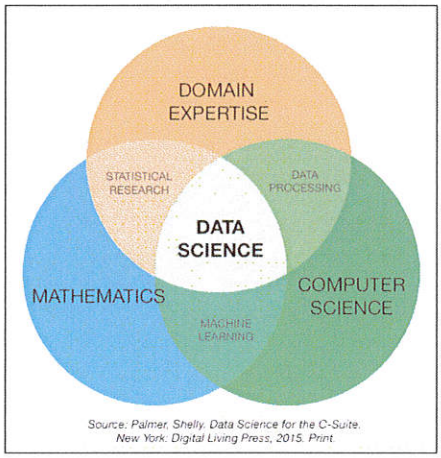
What's Different ?

- Multiple sources generating data much faster than a single computer can utilize (Velocity, Volume)
- Types of data vary. Video, audio, text, etc. (Variability)
- New questions/models are being asked/ employed

Data Science and Analytics (DSA)

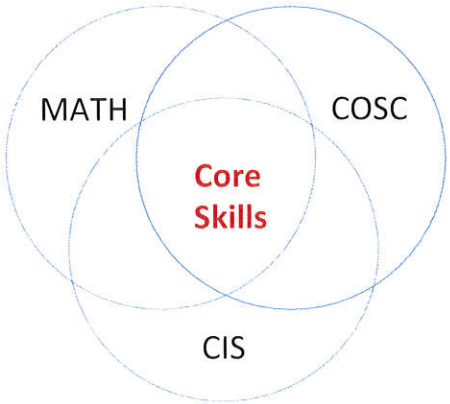


Data Science and Analytics - DSA



Data Science and Analytics (DSA)

DSA at EMU is interdisciplinary!



Data Science and Analytics (DSA)

DSA at EMU is unique !

(stated another way: what's with the A?)

A = Analytics! Translates data-driven insights into decisions and actions. No other undergraduate program does this

Data Science and Analytics (DSA)

DSA at EMU is dynamic!

- Program has been designed to easily incorporate related disciplines (Program Input Document).
- Any discipline is invited to add advanced tracks! Possibilities include:
  - Chemistry – computational chemistry
  - Biology – Genomics/Proteomics
  - Health and Human Services – Healthcare Analytics

Data Science and Analytics (DSA)

DSA at EMU is already here !
Our new Data Scientists:

Roxanne Katus, Ph.D. – Mathematics
Andrey Kashliev , Ph.D. – Computer Science
Yaman Roumani, Ph.D. – Information Systems

Data Science and Analytics (DSA)

The perpetrators:

Stevan Mrdalj – Computer Information Systems
Andrew Ross – Mathematics
William Sverdlik – Computer Science

Data Science and Analytics (DSA)

The abettors:

Chris Gardiner – Mathematics
Augustine Ikeji – Computer Science
Debra Ingram – Mathematics

Data Science and Analytics (DSA)

Questions ?