# An Undergraduate Degree in Data Science: Curriculum and a Decade of Implementation Experience

Paul Anderson, James Bowring, Renée McCauley, George Pothering and Christopher Starr
College of Charleston
Charleston, SC
{ andersonpe2 | bowringj | mccauleyr | potheringg | starrc } @ cofc. edu

## ABSTRACT
We describe Data Science, a four-year undergraduate program in predictive analytics, machine learning, and data mining implemented at the College of Charleston, Charleston, South Carolina, USA. We present a ten-year status report detailing the program's origins, successes, and challenges. Our experience demonstrates that education and training for big data concepts are possible and practical at the undergraduate level. The development of this program parallels the growing demand for finding utility in data sets and streaming data. The curriculum is a seventy-seven credit-hour program that has been successfully implemented in a liberal arts and sciences institution by the faculties of computer science and mathematics.

## Categories and Subject Descriptors
K.3 **Computers & Education**]: Computer and Information Science Education – *computer science education, curriculum*

## General Terms
Algorithms, human factors.

## Keywords
Informatics, curriculum, cognate, analytics, data mining, machine learning, big data, contextualization, interdisciplinary, data science.

## 1. INTRODUCTION
By the late 1980s, the exponential growth in accumulated electronic data generated new interest in statistical machine learning to discover value in the data [15,16]. The new discipline of Knowledge Discovery in Databases (KDD) [16] emerged and took root in graduate programs for data-rich areas such as bioinformatics and drug discovery. The ACM Special Interest Group on Knowledge Discovery and Data Mining (SIGKDD) was established in 1998 to support this new field [1].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
*SIGCSE'14*, March 5–8, 2014, Atlanta, Georgia, USA.
Copyright © 2014 ACM 978-1-4503-2605-6/14/03…$15.00.
DOI string from ACM form confirmation

In 2006, SIGKDD published a two-course curriculum proposal in data mining [11]. In 2011, after 20 years of disciplinary maturation, at least 19 academic programs were available in the US and Canada with a connection to KDD either as a focus or as a supplement to a primary area of study [7]. By August 2013, the number of programs listed had grown to 61 of which only four were clearly undergraduate programs [8] – at the College of Charleston, George Mason, Indiana University and the University of Iowa.

More recently, data science and the related issues surrounding the use and utility of big data formed a cohesive theme for the June 2013 issue of IEEE Computer [13]. Media outlets have also covered stories about big data and emerging programs in higher education [14].

Implementation of the College of Charleston program began in May 2003, when a group of faculty in math, statistics, computer science, and biology believed that it would be possible to create a four-year undergraduate program based on KDD as an emerging discipline. The primary challenge was finding a way to offer the new degree within a liberal arts environment that had fifty-six hours of general education requirements. This paper presents the program development effort, the results of deploying the program, and a review of data collected during the last eight years of deployment.

## 2. CURRICULAR FOUNDATIONS
### 2.1 Origin of the Name and Concept
William Agresti, in an August 2003 article [2], coined the term "discovery informatics" to refer to "the study and practice of employing the full spectrum of computing and analytical science and technology to the singular pursuit of discovering new information by identifying and validating patterns in data." He also suggested that "discovery informatics" is a methodology for identifying and validating patterns in data and that belongs alongside artificial intelligence and simulation in the ACM Computing Classification System (CCS). He connected this new methodology to descriptors and topics in the 1998 classification system by identifying descriptors related to information storage and retrieval, AI, document and text processing, pattern recognition, and database management as parent disciplines of discovery informatics. In 2012, the ACM updated the CCS to include a number of related categories such as "data mining" and "web mining" as well as 109 concepts containing the word "data" [4].

The Agresti article appeared while our committee was working on this new undergraduate curriculum. The similarity between what Agresti defined as "discovery informatics" and the

curriculum that was already in development inspired us to choose Discovery Informatics (DI) as the original name for the new degree program. Coincidentally, and unbeknownst to us at the time, another term, "data science", was coined in 2001 by William S. Cleveland [5]. His focus was on a multi-disciplinary expansion of the field of data analysis. The two terms competed for relevance during the ensuing decade and by 2010 it was clear the term "data science" was ascendant. Therefore, as a strategy to increase the visibility of our program, we moved to rename the degree as Data Science (DS) in 2012.

## 2.2 Knowledge Discovery in Datasets

Research related to KDD in the late 1980s had reached such a volume as to be evidence of an emerging discipline. Workshops in KDD were held at AAAI conferences in 1989, 1991 and 1993. Piatetsky-Shapiro and Frawley published a book on KDD in 1991 [16] that encapsulated then current research in the area. ACM's Special Interest Group on Knowledge Discovery in Data (SIGKDD) was formed and has become the primary professional organization for knowledge discovery. It has held an annual conference since 1995 and has published a biannual academic journal, "SIGKDD Explorations," since 1999. Other professional organizations that play a significant role in the advancement of KDD are the Institute of Electrical and Electronics Engineers (IEEE), the Institute for Operations Research and the Management Sciences (INFORMS) and the Society for Industrial and Applied Mathematics (SIAM). An important on-line source for current news in the data mining/knowledge discovery community is KDnuggets, founded by Piateski-Shapiro in 1997. KDnuggets currently lists over 15 journals in data mining and knowledge discovery. It also shows that in 2013 alone there were 82 business-related or research meetings and conferences in these areas.

Initially KDD entered into the academic curriculum in 2001 when Carnegie Mellon University launched the first Ph.D. program named Computational and Statistical Learning [12]. During the process of developing our DS degree, a member of our committee visited CMU to learn more about their program. The CMU graduate program in KDD was seminal, and it influenced the design of our DS undergraduate degree program. Three years after the design of our program, a working group of the SIGKDD published Version 1 of the Data Mining Curriculum [11].

## 2.3 Computing Curricula

Since 1978, computer science educators have looked to the ACM and the IEEE for guidelines for building undergraduate computer science courses and degrees. The Computer Science volume of Computing Curricula (CC2001) [10] was the most recent set of guidelines available in 2005. CC2001 outlines a set of recommendations for undergraduate degree programs in computer science. It includes a body of knowledge consisting of 132 knowledge units within fourteen core knowledge areas (KAs) of computer science. The KAs are listed in Table 1. For each knowledge unit, it includes multiple learning objectives.

CC2001 was consulted during the development of our DS program. Knowledge units within nine of the fourteen KAs were found to be appropriate to data science. Nearly 200 learning objectives over eight core areas were chosen as fundamental to this new degree. (See Appendix for the KAs and associated learning objectives chosen.) A revision of CC2001 published in 2008 maintained the original fourteen KAs, in which data mining is only an elective course recommendation in the Information Management section.

**Table 1: Knowledge areas specified in CC2001 & CC2008.**

| | |
|---|---|
| Algorithms and Complexity | Information Management |
| Architecture and Organization | Net-Centric Computing |
| Computational Science | Programming Fundamentals |
| Discrete Structures | Programming Languages |
| Graphics and Visual Computing | Operating Systems |
| Human-Computer Interaction | Social and Professional Issues |
| Intelligent Systems | Software Engineering |

A new curriculum volume, CC2013, was recently published in "draft ironman" form [3]. It maintains much of the content from CC2001 & CC2008. It reorganizes and refactors content between several of the KAs, introducing six additional KAs and renaming others. The new KAs are listed in Table 2. The objectives chosen for the DS degree persist in CC2013. Data mining is elective material.

**Table 2: New knowledge areas[1] specified in CC2013.**

| |
|---|
| Networking and Communication |
| Platform-based Development |
| Parallel and Distributed Computing |
| Security and Information Assurance |
| Software Development Fundamentals |
| Systems Fundamentals |

# 3. PROGRAM DESIGN

The design of the DS curriculum presented unique problems because of the need to accommodate several disciplines and incorporate multiple learning objectives.

## 3.1 Interdisciplinary Faculty

DS is an interdisciplinary program that mirrors the interdisciplinary nature of data. Faculty members from mathematics, statistics and computer science, as well as from other disciplines collaborated to design the program. These faculty members routinely use informatics processes and tools in successful research programs such as in bioinformatics and geoinformatics.

## 3.2 Multidisciplinary Learning Objectives

The design of the curriculum began with the specification of learning objectives. These were derived largely from mathematics, statistics, and computer science. CC2001 proved to be a convenient framework on which to build the program. See Appendix for details.

The DS curriculum balances theory and concepts, tools and current techniques, with elements from a range of contributing disciplines, including computer science, artificial intelligence, mathematics, statistics, data mining, communication, and discussions of social and ethical issues associated with DS. The following outline shows the top two levels of the DS curriculum topic list. See Appendix for details.

1. **Large data sets / streams:** create/design, access, clean, analyze, aggregate, organize, visualize

[1] The CC2001 KA titles "Programming Fundamentals" and "Net-Centric Computing" are not part of CC2013. Their content appears in other KAs.

2. **Database:** design, storage, query, modeling
3. **AI techniques:** genetic algorithms, neural networks, Bayesian networks, intelligent agents, machine learning, pattern matching, heuristic search, knowledge representation and ontologies, natural language processing
4. **Software and Algorithms:** design, programming, testing, algorithms and analysis
5. **Information retrieval:** information theory, data mining, text mining, image mining, indexing, content analysis, linguistic processing, abstracting, search and retrieval, information filtering, query formulation
6. **Mathematics:** logic and counting, discrete structures, statistics, linear algebra, modeling and simulation
7. **Oral and written communication:** effective communication
8. **Social, ethical and legal issues:** privacy and security, property, policy, information validation, professionalism

For each topic in the DS curriculum, one or more learning objectives were specified. CC2001 heavily influenced the specification of learning objectives for the computer science content in the curriculum since, in many cases, appropriate learning objectives already existed. For example, for the Machine Learning topic under AI techniques (above), we incorporated six of the eight learning objectives from the CC2001 knowledge unit "IS8: Machine Learning and Neural Networks" into the DS curriculum.

As noted above, three years after the design of our Data Science program, a working group of the SIGKDD published Version 1 of the Data Mining Curriculum [11]. As with DS program design, the Data Mining Curriculum combines foundational and advanced topics, but with a narrow focus on data mining topics. The model sets prerequisites for entry and success in the Data Mining Curriculum. The DS program design is broader in scope than data mining, and includes topical areas having specific learning objectives. It also includes instruction in areas assumed to be prerequisites in the Data Mining Curriculum, including databases, algorithms, data structures, statistics, and linear algebra.

# 4. PROGRAM DETAILS

The design and development of the DS program required blending courses from several academic departments and organizing them as cognates to support integration across the liberal arts and sciences environment.

## 4.1 Multidisciplinary Course Requirements

The program objectives were satisfied by mapping the learning outcomes to undergraduate courses. The result was the use of eleven existing courses and the development of seven new courses. The core requirements are given in Table 3. Courses not previously available and created for the program are in italics. The new courses complete the set of core courses required to deliver the program learning objectives.

Only three of the new courses are unique to the DS program. Four have been incorporated into other degree programs in mathematics and computer science as elective courses and thus enjoy enrollments from two or more student major populations.

**Table 3: Fifteen courses in mathematics, computer science, and data science that form the core of the Data Science program.**

| Core Course Requirements | Semester Credit Hours |
|---|---|
| Data Science | Total: 9 |
| *Introduction to Data Science* | 3 |
| *Dataset Organization & Management* | 3 |
| *Data Science Capstone* | 3 |
| Computer Science | Total: 19 |
| Computer Programming I | 3 |
| Computer Programming II | 3 |
| Computer Programming I Lab | 1 |
| Data Structures & Algorithms | 3 |
| *Advanced Algorithms* | 3 |
| *Data Mining* | 3 |
| Principles of Artificial Intelligence | 3 |
| Mathematics | Total: 26 |
| Introductory Calculus | 4 |
| Linear Algebra | 3 |
| Discrete Structures I | 3 |
| Calculus II | 4 |
| Statistical Methods I | 3 |
| Statistical Methods II | 3 |
| *Statistical Learning I* | 3 |
| *Statistical Learning II* | 3 |

Due to space constraints, only the three courses unique to the DS program are detailed here:

**Introduction to Data Science** This course has no prerequisites, and it provides an overview of the field. Its primary focus is introducing the field to a wide audience. The course focuses on the algorithms and tools that are required of a data scientist. Topics include: knowledge discovery techniques; tools for the analysis of large data sets; the data science process; inductive data-driven modeling.

**Dataset Organization & Management** This course covers the principles of big data organization and management. Topics include: relational and entity relationship models; local and distributed storage and access; preparation and management of datasets for analysis; data cleaning, reorganization and security.

**Data Science Capstone** A project-based capstone course for the application of knowledge discovery and data mining tools and techniques to large data repositories or data streams within the student's cognate area. The capstone experience is designed as an opportunity to synthesize the concepts learned throughout the curriculum to design novel software and algorithms for a domain specific problem related to a student's cognate area.

Details on these and all other required courses are available from the program website: `http://datascience.cofc.edu/`

## 4.2 Cognates for Contextualization

Initially, faculty in disciplines who had data-intensive research programs identified a set of ten cognates from across the liberal arts and sciences. The DS program requires that students complete one cognate. The number of cognates increased to 14 over the first five years of program implementation. Table 4 lists the cognates available in 2013 for students of the DS program.

**Table 4: Cognates that establish a context for study and application of KDD concepts and tools.**

| Cognates | |
|---|---|
| Accounting | Geoinformatics |
| Biomechanics | Molecular Biology |
| Customer Relationship Mgmt. | Organismal Biology |
| e-Commerce | Physics and Astronomy |
| Economics | Psychology |
| Exercise Physiology | Sociology |
| Finance | Supply Chain Mgmt. |

Each cognate requires between fifteen and twenty-two semester credit hours. The set of courses required for these cognates includes introductory and advanced courses in the cognate's domain discipline.

The purpose of a cognate is to introduce a student to an application domain. Courses in a cognate include coverage of domain terminology and concepts, data acquisition devices, common error types in the domain data and processes, and problems typical of the domain. The capstone course defines a custom project for each student to use KDD in the cognate domain under the direction of the cognate director or a faculty member doing data-driven research in that domain.

## 4.3 Integration into the Liberal Arts Environment

At the College of Charleston, degrees in English, philosophy, and history each require 10 courses (thirty semester credit hours), intentionally providing options for students to select courses from other disciplines. The DS program, which requires up to seventy-seven credit hours depending on cognate, leaves little room for free electives.

The difference, however, between DS and the other programs requiring more credit hours than traditional liberal arts degree programs lies in the interdisciplinary nature of KDD. The core courses are equivalent to a liberal arts degree in scope but, additionally, DS exposes students to at least four disciplines (see Tables 3 and 4).

# 5. EXPERIENCES

The program costs have been marginal. Most significantly, no new faculty were needed because out of the 18 courses in the DS core curriculum, eleven were existing courses and four of the seven new courses also contribute to existing undergraduate majors. All courses in the cognates were existing courses.

The number of enrollments continues to grow, but remains small. Although it was always assumed that the rigorous mathematics and computer science courses together with the requirement of a cognate would keep the total enrollment small (never more than 50 majors in an institution with over 10,000 undergraduates) even this level of interest has not yet been reached. Figure 1 gives enrollment and graduation data per year from 2005 to 2013. No estimates for graduation rates were given in the initial proposal, however the institution's latest six-year graduation rate is 66.1%.

The graduates in 2007 and 2008 were already enrolled as computer science majors when the DS program was launched and were able to incorporate the DS requirements into their programs of study. The decline in majors number of majors in 2008 was disappointing, but the poor state of the U.S. economy at that time may have discouraged students from attempting innovative programs when the prospect of employment for the program's graduates was uncertain. It should be noted that, while there have been years of little or no growth, the overall trend has been positive, so there is reason to be optimistic that the number of majors will continue to increase.

One desired outcome of the program, which is being achieved, is to attract strong students. The average GPA, on a 4 point scale, of students who were matriculated in the DS program in Fall 2013 was 3.34 compared to an institutional average of 2.96 [6].

Another positive outcome is that all cognates have been chosen at least once. The most heavily favored cognates have been molecular biology, geoinformatics, psychology, and e-commerce. Anecdotal evidence from department chairs affiliated with some of the cognate areas indicates a strong interest with the DS program and its majors among new faculty candidates.

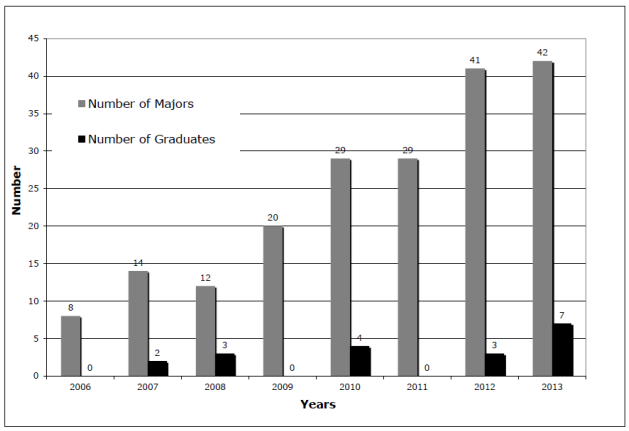



**Figure 1: Anticipated enrollments, actual enrollments and number of graduates in the Data Science program.**

Of those graduates who have not gone directly to graduate school, all have found career placements in data science, informatics, or computer science positions. Nevertheless, some of the graduates have experienced difficulties during their job search. While this could be attributed to economic conditions, some of the difficulty might be due to a lack of name recognition of the program, which we believe to be resolved with the change to Data Science from Discovery Informatics.

Institutionally, the DS program has been beneficial to the computer science and statistics students. Computer science students now have access to regularly-offered courses in data mining, AI (including machine learning), and advanced algorithms. The mathematics department has created a new statistics track, in part due to the two new upper-level courses in statistical learning required in the DS program.

# 6. DISCUSSION AND FUTURE PLANS

In 2003, the computer science department at the College of Charleston offered three undergraduate programs: a BS in Computer Science (ABET accredited), a BA in Computer Science, and a BS in Computer Information Systems (CIS). The CIS program is an interdisciplinary program including computer science and business courses. The DS program, launched in 2005, provided a second interdisciplinary computer science program by integrating computer science, mathematics,

statistics, and a cognate area. The rigor of this interdisciplinary program attracts high-achieving students.

Several benefits of the DS program development process were:

- The institution has a small and stable faculty. This helped develop professional relationships across disciplines that led to the support of the program and the students who engage in it.
- Faculty drove the program development with the support of administration at all levels, including the institution president. This fostered a heightened level of trust and sense of partnership between the faculty and the administration.
- The new degree attracted new enrollments to existing computer science courses.
- The program generated students capable of being undergraduate research assistants in research labs requiring analytics.

We have learned many lessons and encountered several challenges while implementing our undergraduate program in DS that we believe an institution should consider before starting a similar program:

(1) We believe the lack of a community-supported and established name negatively affected program awareness, which led to lower than expected recruitment numbers. After the community converged on the name Data Science (current name of the program), we believe that name recognition is no longer a problem.

(2) Students with one or two calculus courses prior to collegiate study have more flexibility in their overall schedule. This helps the student navigate through the program since DS requires courses from multiple disciplines.

(3) The interdisciplinary nature of the program leads to advising challenges. We have addressed these by having the director of the DS program advise most majors and minors.

(4) Student retention to graduation could be negatively impacted due to the computational and mathematical rigor of the program. To counteract this, we require a high impact introductory course of all DS majors. This breadth-first introduction to the major engages students in the application of data science. For example, we have the students participate in the online data science competition community (www.kaggle.com). It also helps to build a community or cohort of DS students, making students feel a part of the program early on.

We are working on several initiatives to improve the degree. First, we are revisiting our 14 cognate areas to see if they can or should be simplified or modified. This is an ongoing evaluation that we conduct every few years. We are also examining the curriculum to see if there are ways to increase scheduling flexibility. Finally, we continue to refine the learning objectives for the DS curriculum; this will be necessary until a body of knowledge is derived by the KDD community for undergraduate programs in data science.

Our experience proves that a data science degree program can be successfully mounted at the undergraduate level. To do so requires an interdisciplinary team of computer scientists, mathematicians, and statisticians supported by university administration and domain experts in the specific disciplines. The rigor of this interdisciplinary program attracts high-achieving students, which strengthens all of our programs. DS graduates have been successful in industry positions specific to DS, but also those that typically hire CS graduates. Moreover, graduates have pursued graduate study at the Masters level in Analytics, and others have gone on to graduate study in Computer Science.

## 7. ACKNOWLEDGEMENTS

Many faculty members contributed their insights and time to the development of this program They are (alphabetical by last name): Paul Buhler, Computer Science; Alex Kasman, Mathematics; Bill Manaris, Computer Science; Norine Noonan, Dean, School of Sciences and Mathematics; Alan Strand, Biology; and James Young, Mathematics. Many other faculty members participated in the program approval and implementation process. Thanks to them all.

## REFERENCES

[1] ACM SIGKDD Charter. SIGKDD, June 10, 1998. Accessed September 6, 2013. http://www.acm.org/sigs/bylaws-contents/kdd_bylaws/

[2] Agresti, W.W., Discovery Informatics, *Communications of the ACM*, August 2003, Vol. 46, No. 8, pp 25-28.

[3] CC2013 Ironman. The Joint Task Force on Computing Curricula. ACM, IEEE-Computer Society. Accessed September 5, 2013. http://ai.stanford.edu/users/sahami/CS2013/ironman-draft/cs2013-ironman-v1.0.pdf

[4] CCS2012. Accessed September 6, 2013. http://dl.acm.org/ccs_flat.cfm

[5] Cleveland, W.S., Data Science: An Action Plan for Expanding the Technical Areas of the Field of Statistics, ISI Review, 69, 2001, 21-26.

[6] College of Charleston 2013 Planning and Reference Guide. Accessed September 5. 2013. http://irp.cofc.edu/facts-and-publications/PRFB-Main-Menu.php

[7] Data Mining Education. KDnuggets. http://www.kdnuggets.com/education/usa-canada.html. Accessed January 11, 2011.

[8] Data Mining Education. KDnuggets. http://www.kdnuggets.com/education/usa-canada.html. Accessed August 28, 2013.

[9] Fayyad, U.M., and Uthurusamy, R. editors, Knowledge Discovery in Databases: Papers from the 1994 AAAI Workshop, Seattle, Washington, July 1994. AAAI Press, Technical Report WS-94-03, ISBN 0-929280-73-3, 473 pp.

[10] The IEEE-ACM Joint Task Force on Computing Curricula. Computing Curricula 2001: Computer Science, December 15, 2001.

[11] Intensive Working Group of ACM SIGKDD Curriculum Committee. Data Mining Curriculum: A Proposal (Version 1.0), April 30, 2006. www.sigkdd.org /curriculum.php.

[12] Machine Learning Ph.D. Programs, Machine Learning Department, Carnegie Mellon University, http://www.ml.cmu.edu/prospective_students/phd_programs.html

[13] Michael, M. and Miller, K.W., guest ed. Big Data: New Opportunities and Challenges. IEEE Computer. June 2013.

[14] Miller, C.C., Data Science: The Numbers of Our Lives. *New York Times*. April 11, 2013.

[15] Piatetsky-Shapiro, G. Knowledge discovery in real databases: A report on the IJCAI-89 Workshop. AI Mag. 11, 5 (January 1991, 68-70.

[16] Piateski-Shapiro, G., and Frawley, W., Knowledge Discovery in Databases. MIT Press, Cambridge, MA. 1991.

# APPENDIX

**Topic List and Learning Objectives[2] for Undergraduate Program in Data Science**

| | |
|---|---|
| **1.** | **Large data sets / streams - Creation/design/; Access; Clean; Analyze; Aggregate; Organize; Visualization** |
| | a. Information Management, IM4, Relational databases, objectives 1-5<br>**b.** Graphics and visualization**,** GV9, Visualization**,** objectives 1-4 |
| **2.** | **Database - Design; Storage; Query; Modeling** |
| | **a.** Information Management, IM 6, Relational database design, objectives 1-7<br>**b.** Information Management, IM 8, Distributed databases, objectives 1-5<br>**c.** Information Management, IM 5, Database query languages, objectives 1-5<br>**d.** Information Management, IM 3, Data modeling, objectives 1-6 |
| **3.** | **AI techniques - Genetic algorithms; Neural networks; Bayesian networks; Intelligent agents; Machine learning; Pattern matching; Heuristic search; Knowledge representation and ontologies; Natural language processing** |
| | a. Intelligent Systems, IS8. Machine learning and neural networks, objectives 1-8<br>b. Intelligent Systems, IS6. Agents objectives 1-6<br>c. Intelligent Systems, IS2, Search and constraint satisfaction, objectives 1-7<br>d. Intelligent Systems, IS3, Knowledge representation and reasoning, objectives 1-4<br>e. Intelligent Systems, IS5, Advanced knowledge representation and reasoning, objectives 1-6<br>f. Intelligent Systems, IS7, Natural language processing, objectives 1-5 |
| **4.** | **Software and Algorithms - Design; Programming; Testing; Algorithms and analysis** |
| | a. Programming Fundamentals, PF5, Event-driven programming, objectives 1-3<br>b. Software Engineering, SE9, Component-based computing, objectives 1-5<br>c. Programming Fundamentals, PF2, Algorithms and problem-solving, objectives 1-5<br>d. Programming Fundamentals, PF3, Fundamental data structures, objectives 1-8<br>e. Programming Fundamentals, PF1. Fundamental programming constructs, objectives 1-6<br>f. Algorithms and Complexity, AL1. Basic algorithmic analysis, objectives 1-5<br>g. Algorithms and Complexity, AL2. Algorithmic strategies, objectives 1-8<br>h. Algorithms and Complexity, AL3. Fundamental computing algorithms, objectives 1-7<br>i. Algorithms and Complexity, AL5. Basic computability, objectives 4 & 5<br>j. Algorithms and Complexity, AL6. The complexity classes P and NP, objectives 1-3<br>k. Algorithms and Complexity, AL8, Advanced algorithmic analysis, objectives, 1-4 |
| **5.** | **Information retrieval - Information theory; Data mining; Text mining; Image mining; Index** |
| | a. Information Management, IM 11, Information storage and retrieval, objectives 1-5<br>b. Information Management, IM 10, Data mining, objectives 1-8<br>c. Information Management, IM 12, Hypertext and hypermedia, objectives 1-5<br>d. Information Management, IM 13, Multimedia information and systems, objectives 1-5 |
| **6.** | **Mathematics - Logic and counting; Discrete structures; Statistics; Linear Algebra; Modeling and simulation** |
| | a. Discrete Structures, DS2, Basic logic, objectives 1-4<br>b. Discrete Structures, DS4. Basics of counting, objectives<br>c. Discrete Structures, DS1. Functions, relations, and sets, objectives, 1-4<br>d. Discrete Structures, DS3. Proof techniques, objectives 1-4<br>e. Discrete Structures, DS5. Graphs and trees, objectives 1-4<br>f. Discrete Structures, DS5. Discrete probability, objectives 1-4<br>g. Computational Science and Numerical Methods, CN3, Modeling and simulation, objectives 1-4 |
| **7.** | **Oral & written communication** |
| **8.** | **Social, ethical and legal issues - Privacy and security; Property; Policy; Information validation; Professionalism** |
| | a. Social and Professional Issues, SP7. Privacy and civil liberties, objectives 1-5<br>b. Social and Professional Issues, SP6. Intellectual property, objectives 1-5<br>c. Social and Professional Issues, SP4, Professional and ethical responsibilities, objectives 1-6<br>d. Social and Professional Issues, SP5, Risks and liabilities of computer-based systems, objectives 1-4<br>e. Social and Professional Issues, SP2. Social context of computing, objectives 1-5 |

[2] All learning outcomes are from CC2001 [10]