

5-2019

Designing a general education course on the societal impacts of artificial intelligence

Vincent Rollins

University of Tennessee, Chattanooga, pgn961@mocs.utc.edu

Follow this and additional works at: <https://scholar.utc.edu/honors-theses>

Part of the [Artificial Intelligence and Robotics Commons](#)

Recommended Citation

Rollins, Vincent, "Designing a general education course on the societal impacts of artificial intelligence" (2019). *Honors Theses*.

This Theses is brought to you for free and open access by the Student Research, Creative Works, and Publications at UTC Scholar. It has been accepted for inclusion in Honors Theses by an authorized administrator of UTC Scholar. For more information, please contact scholar@utc.edu.

Designing a General Education Course on the Societal Impacts of Artificial Intelligence

Vincent Michael Rollins

Honors College Thesis
The University of Tennessee at Chattanooga

Examination Date: April 19, 2018

Dr. Gregory O'Dea
UC Foundation Professor and
Associate Dean of the Honors College
Thesis Director

Dr. Matt Matthews
Department of Mathematics Professor and
Interim Vice Provost
Department Examiner

Contents

Introduction	2
Honors College Brock Scholars Course Proposal	4
Syllabus	8
Course Outline	11
Week 1	13
Week 2	15
Week 3	24
Week 4	26
Week 5	27
Week 6	30
Week 7	33
Week 8	34
Weeks 9 & 10	35
Week 11	38
Week 12	39
Weeks 13 & 14	41
Week 15	43
Conclusion	44
References	46

Introduction

Nearly every piece of technology we use today was invented on the basis of a single question: “How can a particular job or process be made easier?” The moment we stop asking ourselves this question is the moment when technology stops progressing. Our perseverance in this endeavor eventually revealed that a process is easiest when it is automated by something we create. What, then, would happen if our process of invention itself were automated? Can we create a machine that can learn and think for us? This is the entire idea behind creating artificial intelligence, otherwise known as AI.

Research in AI began in the 1960’s and has progressed exponentially, with every day bringing a new development that moves us closer and closer to an artificially intelligent program. However, the potentially widespread implications of this technology are often underrepresented to the general public. Even more than the internet, AI has the potential to change every aspect of our world, but as with the early days of the World Wide Web, relatively few people are aware of that potential. This thesis proposes to answer the following questions, with the intent of incorporating the answers into a general education course syllabus: What are the most important concepts in understanding the importance of this emerging technology? Where are we now in our endeavors, and what are the predictions moving forward? What are the economic, philosophical, cultural, and ethical implications of AI? Would artificial intelligence be considered living/conscious, and if so, what would that imply? Questions such as these help direct the design of a course that informs students of the emergence of artificial intelligence and its truly extensive repercussions.

Most colleges, including UTC, already offer an artificial intelligence course (CPSC 4440) as part of their computer science curricula. Such courses are meant to explain the technology

behind these elaborate systems, but these courses often neglect extensive coverage of the real-world impacts of the technology itself. UTC also offers a course entitled “Ethical and Social Issues in Computing” that does convey the importance behind the advances of computer technology and its impacts, but this course is practically available only to computer science majors. There is no generalized and widely available course that covers the technological, economic, cultural, philosophical/theological, and ethical concerns that come with the implementation of artificial intelligence.

My main endeavor in this thesis has been to review the literature of the fields involved and decide what collection of reading/film assignments best covers the necessary information when it comes to the widespread impacts of AI. I then constructed a course syllabus with clearly defined assignments and a schedule of topics. Furthermore, I discuss each decision in the construction of the syllabus and explain how and why each topic will be addressed when the course is offered in the fall semester of 2018. The final product of my research is this thesis, detailing the considerations that went into the design of the course; every anticipated discussion topic, reading, and writing assignment is laid out in detail, and the decisions for every area of the class are explained. A final syllabus for the course has been drafted as well as a completed course proposal form containing clear explanations of how this course meets the established purpose and learning outcomes of UTC’s Thought, Values, and Beliefs general education category.

Honors College Brock Scholars Course Proposal

Name(s) and department(s) of faculty member(s) submitting the proposal:

Name: Vincent Rollins & Dr. Greg O’Dea

Department: Computer Science &
University Honors

Email: pgn961@mocs.utc.edu

Phone: 423-310-6767

Title of course: The Impacts of Artificial Intelligence

Indicate preference for semester to be taught:

Fall 2018

Spring 2019

Have you taught the course you are proposing in the Brock Scholars Program before? If so, please indicate when you last offered the course.

N/A

Please indicate the General Education category in which you intend to offer the course:

Historical Understanding

Natural Sciences (Non-lab)

Literature

Natural Sciences (Lab)

Thought, Values, and Beliefs

Mathematics

Visual and Performing Arts

Statistics

Behavioral and Social Sciences

Non-Western Cultures

Course Description:

Please include as much information as possible about the general themes of the course, topics to be covered, and goals to be accomplished in the course.

Research in AI began in the 1960’s and has progressed exponentially, with every day bringing a new development that moves us closer and closer to an artificially intelligent program. However, the potentially widespread implications of this technology are often rendered irrelevant to the general public by means of other exciting news stories. Even more than the internet, AI has the potential to change every aspect of our world, but as with the early days of the World Wide Web, relatively few people are aware of that potential. This course examines the development of artificial intelligence in the modern world and the impacts this emerging technology have had and will continue to have across our culture, including areas such as technology, economics, law, social equity, philosophy, religion, politics, ethics, and literature/movies.

Goal: This course will teach students how to recognize modern AI technologies, analyze current and future impacts of those technologies, and contemplate the cultural advancements and pitfalls of those impacts.

Objective: Upon completion of this course, students should be able to understand both the history of AI and how it relates to the current direction of development, recognize the potential for widespread impact that artificial intelligence offers, comprehend where and how different areas of culture are driving AI development, and analyze how our culture can be better prepared in areas such as job displacement, law precedence, social equity, and cultural mindsets for the emergence of synthetic intellects and electronic entities.

Course Outline:

How do you plan to structure the course? Please provide a brief outline of the course (this may be a weekly schedule of topics, a sequence of "units," etc.).

Weeks 1–2: Intro to AI:

Week 3: Automation and Economics

Week 4: Automation and Education

Week 5: AI and Law

Week 6: Automation and Social Equity; Solutions

Week 7: Big Data and Machine Learning

Week 8: Midterm and Introduction to Part II of the Course

Week 9: Consciousness and AI: philosophy of consciousness and its potential to be programmed

Week 10: Consciousness and AI: history of anthropomorphism and how we define the mind

Week 11: Consciousness and AI: responsibilities as creators and programming ethics into machines

Week 12: Consciousness and AI: what will society look like with another form of intelligence/consciousness?

Week 13: Culture and AI: Current opinions and politics surrounding AI

Week 14: Culture and AI: Popular culture's portrayal of AI

Reading Assignments:

Honors courses should emphasize core use of primary source reading, not textbook learning. Please provide a tentative reading list for your proposed course.

- *Humans Need Not Apply: A Guide to Wealth and Work in the Age of Artificial Intelligence* by Jerry Kaplan
- *Frankenstein* by Mary Shelley
- *I, Robot* by Isaac Asimov
- Excerpts from several other books, including:
 - *The Master Algorithm* by Pedro Domingos
 - *In Our Own Image: Savior or Destroyer? The History and Future of Artificial Intelligence* by George Zarkadakis
 - *Artificial Intelligence: What Everyone Needs to Know* by Jerry Kaplan.
 - As well as several articles and short videos

Written Assignments:

Honors courses are should be writing intensive wherever appropriate to the learning outcomes of the course. Please estimate the amount and the nature of written work that will be expected of students in this course. (Short/long papers? Term papers? Essay exams? Researched writing?)

- 5 Minor writing assignments (~1 page each)
- 1 Film response (3 pages)
- A final response regarding how their personal viewpoint on AI has changed throughout the course

Other Assignments:

Beyond reading and writing, what other kinds of assignments do you envision for the course?

- Vocabulary test
- Midterm Test
- Solving a hypothetical court case involving AI, researching previous court cases to gather precedence

Classroom Procedures:

Honors courses are intended to emphasize active as opposed to passive learning. Will this be primarily a lecture or discussion course? If the course is intended to have significant lecture content, how will interaction be fostered among students and faculty? What specific methods will be used to encourage an active learning experience?

The first half of the course will involve a fair amount of lectures in order to explain many of the concepts of AI. To balance this, the second half of the course will be less grounded and more theoretical. Readings, including works of fiction, will be assigned throughout the course that will be discussed in class. Several assignments will also involve debates amongst the class.

Repeat Offering

If you are proposing a course you have previously taught in Honors, please indicate how you have responded to student comments and evaluations to craft your approach this time around – what have you maintained and changed in this proposed version of the course and why have you made these choices?

N/A

Each faculty member submitting this proposal should have his or her chairperson sign below.

Faculty Signature

Department Head signature

Department Head's signature verifies that, if selected, the faculty member will be released to teach the proposed course sometime in the 2017–2018 academic year

Date submitted: _____

Syllabus

UHON 3530: The Impacts of Artificial Intelligence

Spring 2018

Dr. Greg O'Dea and Vincent Rollins

Office: Honors House, phone: , Email: Gregory-Odea@utc.edu, PGN961@mocs.utc.edu

When/Where: T/R 1:40-2:55 PM. Office hours: TBD.

Lecture Schedule Type: Conventional Methodology Instructional Method

3.000 Credits

Syllabus

Catalogue Description: Change

The goal of this subcategory is for students to engage analytically and critically at least one body of philosophical, ethical, and/or religious thought. Students will examine ways in which such thought can shape decisions and actions. Students are encouraged to check the specific requirements for their major for this category.

Upon completion of the required credit hours in this subcategory, students will be able to:

- Identify the key components of at least one body of thought.
- Explain and analyze a body of thought.
- Apply the unique perspective of the body of thought to a specific problem or question.
- Effectively articulate in writing their individual perspective in relation to the body of thought.

Course Description

Research in AI began in the 1960's and has progressed exponentially, with every day bringing a new development that moves us closer and closer to a truly artificially intelligent program. However, the potentially widespread implications of this technology are often rendered irrelevant to the general public by means of other exciting news stories. Even more than the internet, AI has the potential to change every aspect of our world, but as with the early days of the World Wide Web, relatively few people are aware of that potential. This course examines the development of artificial intelligence in the modern world and the impacts this emerging technology have had and will continue to have across our culture, including areas such as technology, economics, law, social equity, philosophy, religion, politics, ethics, and literature/movies.

General Education and Course Outcomes

This course fulfills a General Education requirement in the Thoughts, Values, and Beliefs category. Upon completion of this course, students should be able to understand both the history of AI and how it relates to the current direction of development, recognize the potential for widespread impact that artificial intelligence offers,

comprehend where and how different areas of culture are driving AI development, and analyze how our culture can be better prepared in areas such as job displacement, law precedence, social equity, and cultural mindsets for the emergence of synthetic intellects and electronic entities.

Required Text

- *Humans Need Not Apply – A Guide to Wealth and Work in the Age of Artificial Intelligence*, by Jerry Kaplan. ISBN: 978-0300213553
- *Frankenstein*, by Mary Shelley. ISBN: 978-0486282114
- *I, Robot*, by Isaac Asimov. ISBN: 978-0553294385
- *In Our Own Image: Savior or Destroyer? The History and Future of Artificial Intelligence*, By George Zarkadakis. ISBN: 978-1605989648

Recommended Text

- *Artificial Intelligence: What Everyone Needs to Know*, By Jerry Kaplan (this book is just a well put together collection of questions and answers that you can refer to throughout the course) ISBN: 978-0190602390

Evaluation and Assessment

Grades will come from quizzes on the readings, tests, essays, and participation in class and will be weighted as such:

Participation: 5%

Quizzes: 10% (only from top 10 quiz grades)

Short Assignments: 25% (only from top 5 assignment grades)

Essays: 30%

Final Presentation: 30%

Grading Scale

All quizzes, assignments, and essays will be graded on a standard 100-point scale. Letter grades receive the following numeric equivalents:

A	B	C	D	F
100-90	89-80	79-70	69-60	>60

Attendance and Participation

Barring any medical emergency or death in the family, you are expected to be present in class and take part in any discussions that arise in class. You are allowed two unexcused absences throughout the semester, after which 5 points will be deducted from your final grade for every class missed beyond those two days. If you have any questions about what might be considered an excused absence then please notify me beforehand and we will discuss. Also, we will often discuss what everyone wrote about in their mini essay assignments the class period after they are turned in, so be prepared to share.

Quizzes

There will be a 5-question multiple choice reading quiz at the start of most classes based on the assigned reading. These will not be difficult and should only take a couple of minutes to complete. There will be around 12-15 quizzes across the semester and only your top 10 quiz grades will be used when calculating your final grade.

Assignments

There will be six assignments inside and outside of class with only your top five grades taken into account. Each assignment is worth 5% and will total up to 25% of your final grade. More information about each assignment will be given when they are assigned.

Essays

There will be three 3-5 page essays throughout the semester, each worth 10%, totaling up to 30% of your final grade. Essays will be turned in to UTC Learn. More information about each essay will be given when they are assigned.

Final Presentation

Instead of a final exam, each student will put together a presentation on how artificial intelligence will affect them as they move into their careers after college. You should plan on using everything you've learned throughout the semester to explain the type of AI that you are discussing as well as being able to explain how it works and in what ways it will affect your career.

Makeup Quizzes, Assignments and Essays

Reading Quizzes will be held at the start of most classes and cannot be made up for unexcused absences. If you wish to make up a quiz for an excused absence then you can come to office hours and do so. Short assignments will be turned into UTC Learn (unless otherwise specified) and your grade will decrease by one letter grade for every day it is late. Essays will always be turned into UTC Learn and will also decrease by one letter grade for every day that it is late.

Communication outside of class

UHON 3530 will use Blackboard (UTC Learn) to post grades, distribute course materials, and communicate outside of class. To enhance student services, the University uses your UTC email address for all communications. Please check your UTC email on a regular basis. If you have problems with accessing your UTC email account, contact the Call Center at 423/425-4000.

Technology Requirements

Students should be able to use Word, the learning management system (UTC Learn), MOCSNet email, and be able to submit files to UTC Learn.

Course Outline

Week #	Topic	Reading	Assignments and Essays	Objective
Week 1	Course Introduction and Establishment of Expectations. Intro to AI: <ul style="list-style-type: none"> ● What is AI? ● Strong vs. Weak AI 	<i>Wait but Why?</i> AI Revolution article	Assignment #1: Find something you're interested in and report on how Artificial Intelligence will impact it. <i>BlackBoard submission and discuss in class</i>	Students should understand the overall importance of artificial intelligence, and be able to recognize the difference between "weak" AI and "strong" AI (also known as "narrow" AI and "general" AI) as they pertain to contemporary and future developments.
Week 2	Intro to AI cont. <ul style="list-style-type: none"> ● History ● Vocabulary 	List of vocabulary words and their definitions and explanations <i>Humans Need Not Apply:</i> Introduction (pages 3-30)	Assignment #2: Vocabulary test. <i>Taken in class</i>	Students should understand the history of AI development and how previous research endeavors in relation to economic incentives have brought us to where we are today. They should also better understand the vocabulary surrounding the technology of artificial intelligence; this includes terms such as: Algorithm, Automation, Machine Learning, Neural Network, etc. And be able to recognize how each relates to the now fluid definition of artificial intelligence.
Week 3	Automation <ul style="list-style-type: none"> ● Economics 	<i>Humans Need Not Apply:</i> Chapters 3 & 4 (p. 51-74) Video: CGP Grey's <i>Humans Need Not Apply</i> YouTube Video (in class)		Students should be able to understand the rapid impact that narrow AI will have on our unprepared job markets, taking special note as to how it pertains to them individually as they transition into their own careers. They should also be able to examine potential solutions to the expected turmoil in that this accelerated job displacement will cause in the equities of society. This will transition smoothly into the current issues and potential solutions involving education reform.
Week 4	Automation cont. <ul style="list-style-type: none"> ● Education 	<i>Humans Need Not Apply:</i> Chapter 8 (p. 131-158)	Assignment #3: Find something about your career where automation might threaten your ability to get a job. <i>BlackBoard submission</i>	Students should be able to recognize the costs and benefits of the current education system and how the rapid development of automation will influence any changes made to our education system. They should also be able to analyze why potential solutions might help improve the situation when it comes to working with automation instead of against it.
Week 5	Automation cont. <ul style="list-style-type: none"> ● Law 	<i>Humans Need Not Apply:</i> Chapter 5 (p. 79-92)	Essay #1: Solve a hypothetical court case involving AI/Automation.	First, students should be able to recognize the impacts that automation will have on how the world of law operates, including how lawyers will come utilize it as well as how it may come to be used in the court itself. Then, students should be able to analyze a hypothetical court of law situation involving an electronic entity and be able to come to a "verdict" themselves using at least one real court case as precedence. Following this, students should be able to recognize the importance that legal precedence will have on such a powerful technology.

Week 6	Automation cont. <ul style="list-style-type: none"> ● Social Equity ● Solutions 	<i>Humans Need Not Apply:</i> Chapters 7, 9, & "Outroduction" (p. 109-127, 161-208)		Students should be able to recognize how the current automation technologies are impacting social equity by looking at the amount a large corporation makes per employee and comparing it to that company's average employee salary/wage. Students should then be able to deduce the reasons behind particular ideas from proposed solutions to the equity crisis that automation could bring on.
Week 7	Big Data and Machine Learning <ul style="list-style-type: none"> ● Intro to schools of thought on AI development 	<i>The Master Algorithm:</i> Prologue		Students should be able to recall the major schools of thought on how AI is being approached from a developmental aspect. Students should then be able to recognize what school of thought is being utilized in a given example of research and development on artificial intelligence.
Week 8	Review of part I and introduction to part II			
Week 9	What is Consciousness? <ul style="list-style-type: none"> ● History of Philosophy of Consciousness ● Can it be programmed? ● What happens then? 		Assignment #4: Prior to starting this half of the course, mini essay on what the student thinks it means to be conscious, intelligent, and alive. <i>BlackBoard submission</i>	Students should be able to recognize the role of consciousness in the past and future of artificial intelligence. They should also be able to recognize how the philosophy of consciousness and its history has influenced the perception of AI in comparison to ourselves. Students should also be able to recognize the difficulties faced in trying to program consciousness (including the difficulties faced in testing for its existence in other beings) as well as what the possible ramifications we might see should it ever be accomplished.
Week 10	Consciousness and AI <ul style="list-style-type: none"> ● History of anthropomorphism ● History of how we define the mind 	<i>In Our Own Image: Savior or Destroyer?</i>	Assignment #5: Mini essay; choose one of the five major paradigm shifts covered in class and come up with a metaphor of your own for the mind-body problem based on the technology of the time period in question. <i>BlackBoard submission</i>	Students should be able to relate our history of consciousness-philosophy to our long history of anthropomorphism in relation to contemporary technology. Students should also be able to recall the major movements throughout history in how we have defined our mind as well as how the technology of the time heavily influence our definition of ourselves.
Week 11	Consciousness and AI <ul style="list-style-type: none"> ● Responsibilities as creators ● Programming ethics into a machine 	<i>Frankenstein</i>	Assignment #6: Mini essay on personal viewpoint in regards to what Frankenstein has to tell us about the responsibility we have when it comes to developing powerful technology. <i>BlackBoard submission</i>	Students should be able to recognize the current stances on what precautions and safeguards should be taken when developing AI. Students should be able to make an argument regarding what our responsibilities as creators should be if we continue to pursue the development of artificial consciousness. This argument can and should utilize a blend of contemporary technology and arguments as well as pieces of science fiction that portray the dangers we pose to ourselves when technology gets out of our hands
Week 12	Consciousness and AI <ul style="list-style-type: none"> ● What does society look like with another form of intelligence/consciousness? 	<i>I, Robot</i> by Isaac Asimov Stories to read: "Robbie," "Runaround," "Reason," "Liar" (p. 1-81, 111-135) "Evidence"	Essay #2: Comparative essay between I, Robot and Frankenstein	Students should be able analyze the potential impacts on society should another form of intelligence/consciousness begin to operate alongside ourselves. Things that should especially be noted are laws and regulations, societal norms, and the interactions that might take place.

		"The Evitable Conflict" (p.206-272)		
Week 13	Today's viewpoints on AI <ul style="list-style-type: none"> Political? Cultural? Where do people stand? Where do you stand now? 	Individually decide on a movie displaying AI as a main theme. A list of options will be provided.		Students should be able to recognize contemporary viewpoints on artificial intelligence in society, including political, cultural, economic, and philosophical. Students should also be able to look back and see how their own viewpoints have changed throughout this course. They should then be able to come to their own conclusions as to if enough is being done to inform the public (as well as each of these individual areas) on the importance of AI.
Week 14	Popular Culture and AI <ul style="list-style-type: none"> Movies/shows, books, games, etc. 		Essay #3: Write a report on how AI was displayed in the movie you chose. What about it was accurate/inaccurate? How did the culture react to that form of AI? Etc.	Students should be able to analyze a particular piece of popular culture centered around artificial intelligence and be able to report on its accuracy in its portrayal of the technology; they should also be able to analyze what the intent of the piece is when it comes to AI.
Week 15	Overview/Review		Final Presentation: Utilizing everything you've learned this semester, give an individualized presentation expanding on how artificial intelligence will affect your career and life as you move past college and into the world.	

Week 1

Week #	Topic	Reading	Assignments and Essays	Weekly Objective
Week 1	Course Introduction and Establishment of Expectations. Intro to AI: <ul style="list-style-type: none"> What is AI? Strong vs. Weak AI 	<i>Wait but Why?</i> AI Revolution article	Assignment #1: Find something you're interested in and report on how Artificial Intelligence will impact it. <i>BlackBoard submission and discuss in class</i>	Students should understand the overall importance of artificial intelligence, and be able to recognize the difference between "weak" AI and "strong" AI (also known as "narrow" AI and "general" AI) as they pertain to contemporary and future developments.

The first week of class must be an introduction to the broad topic of artificial intelligence; basic concepts, terms, and ideas must be established so that the course may expound on itself throughout the rest of the semester. Rather than spending an entire class period trying to lay a groundwork of knowledge for the class, students should be assigned some form of reading that

gives a brief overview of some of the main concepts of artificial intelligence. Also, because this could be the first exposure of AI that a student might come across, the reading must be accessible; it must be something casual, yet informative. Most textbooks and "Introduction to AI" books don't fit these requirements because neither of the two are succinct enough and they run a risk of expecting the reader to have a certain level of knowledge regarding computer science. Therefore, the first reading assignment should come from a source that would typically already have a wide audience—an article from a website of some sort. From my research, the best article that meets all of these requirements would be *The AI Revolution: The Road to Superintelligence* by Tim Urban, from the website WaitButWhy.com. Urban lays out everything we need to know about what artificial intelligence as a technology means to us in the first part of his article, including setting the stage for how impactful AI will be. One of the main topics that Urban covers is the idea of strong AI vs weak AI, giving examples to each. Weak AI is what we see with technology today: specialized programs that can do one or two things really well, while strong AI is what many computer engineers are working towards for the future: generalized programs that can do many things at a level comparable to humans. Establishing this distinction allows for a better understanding of why the term "AI" gets thrown around so much in today's world. Once this is explained, the first assignment will require students to pick anything that interests them and then research how AI will impact that topic as the technology develops. The purpose of this assignment is to show students that AI can and will be applied to almost every aspect of our lives. It is important to illustrate to students early on in the semester that AI technology will impact something that they are interested in, creating a personal level of investment for each student and giving them their own reference point as they continue to learn about artificial intelligence. As soon as all of this groundwork is laid out for the class, the

transition into the rest of the semester will be much smoother. The next week will look at all of the popular terms surrounding artificial intelligence.

Week 2

Week #	Topic	Reading	Assignments and Essays	Weekly Objective
Week 2	Intro to AI cont. <ul style="list-style-type: none"> History Vocabulary 	List of vocabulary words and their definitions and explanations <i>Humans Need Not Apply:</i> Introduction and Chapters 1 & 2 (pages 3-48)	Assignment #2: Vocabulary test. <i>Taken in class</i>	Students should understand the history of AI development and how previous research endeavors in relation to economic incentives have brought us to where we are today. They should also better understand the vocabulary surrounding the technology of artificial intelligence; this includes terms such as: Algorithm, Automation, Machine Learning, Neural Network, etc. And be able to recognize how each relates to the now fluid definition of artificial intelligence.

The second week of class stands apart from the introductory week as being the best opportunity to delve deeper into the technical aspect of artificial intelligence. However, before moving too far into any one topic, it is necessary to lay out basic terms regarding AI that might appear throughout the semester in the readings so as to prevent any misunderstanding that would arise from the frequent use of jargon that comes with any well-established field. Here are the terms that will be used the most throughout this course as well as why students would need to know these terms:

- **Algorithm:** A process or set of rules to be followed in calculations or other problem-solving operations, especially by a computer.

(Oxford Dictionaries, Algorithm)

While students won't need to know much about computer science and programming throughout this course, they will at least need to understand the concept of algorithms. No matter how complicated the computer programs that will be discussed in class, it's important for students to keep in mind that all computer processes can be broken down into logic. Promoting this mindset of simplification throughout the course will help prevent students from being

discouraged when faced with overwhelming explanations of computing processes; a computer is just following an algorithm, no matter how complex.

- **Black Box:** Anything that has mysterious or unknown internal functions or mechanisms.

(Merriam-Webster, Black Box)

Of all the ideas that will be covered throughout this course, the concept of a "black box" is by far the most important. It is imperative that every student recognize their individual experiences of attributing concepts and ideas that they don't quite understand to a black box. However, it's also important to note that the usage of black boxes isn't always a bad thing--it's quite useful not having to know everything about every tool we use on a daily basis. It is apparent that most people in society are operating with a black box around advancements in technology, especially artificial intelligence, and as such are misunderstanding it and its capabilities.

- **Learning:** "A process of performing temporally sequential generalizations, by taking prior experience into account in future analyses, just as reasoning by analogy is a matter of using knowledge from one domain as a novel context with which to generalize about another."

(Kaplan, *Artificial Intelligence: What Everyone Needs to Know*, p. 6)

Another one of the broad concepts that must be covered when talking about artificial intelligence is how we define and understand "learning" as it pertains to computers and how we compare that to our own intelligence and learning processes. I use Kaplan's definition because it encompasses both the abilities that we have as humans as well as the potential abilities that artificial intelligence might showcase as technology advances.

- **Hardware:** The machines, wiring, and other physical components of a computer or other electronic system.

(Oxford Dictionaries, Hardware)

- **Software:** The programs and other operating information used by a computer.

(Oxford Dictionaries, Software)

These two terms are necessary to explain the basic operating systems that are utilized by computers and technology in general. However, the distinction between hardware and software will become more important as topics involving robotics are discussed later on in the course.

- **Machine Learning:** A field of computer science that gives computer systems the ability to "learn" (i.e., progressively improve performance on a specific task) with data, without being explicitly programmed.

(Wikipedia, Machine Learning)

Moving on to terms that are more specific to the study of artificial intelligence, machine learning is one of the major fields of AI that will be discussed throughout the course and this definition is meant to serve as a brief introduction to the concept. And while this topic will be covered more in depth in week 7, it must also be explained early on because it plays heavily in the structure of automation.

- **Big Data:** Extremely large data sets that may be analyzed computationally to reveal patterns, trends, and associations, especially relating to human behavior and interactions.

(Oxford Dictionaries, Big Data)

If machine learning programs are the engines driving the progress of automation, then Big Data is the fuel for those engines. In the Information Age, data is being collected on every aspect of our lives, and while this course will eventually look into the effects of mass data collection on the general public, it's important to explain this term early on in the course because it may be referenced in the conversation of machine learning.

- **Artificial Neural Networks:** A machine learning technique that utilizes interconnected layers of neurons, each acting as its own low-level function. After proper training, a neural network breaks down an overall input into smaller parts and analyzes the individual parts in relation to each other in order to come to a conclusion.

(Kaplan, *Artificial Intelligence: What Everyone Needs to Know*, p. 28-29)

Neural Networking is considered one of the most popular subfields of machine learning and is often used imprecisely in news headlines. It will be the primary machine learning technique that this course will use as an example to explain how a machine learning algorithm simply takes in an input and produces an output. To supplement this definition, students will be walked through a simple step-by-step example of how a neural network "learns" and how it makes a decision.

- **Supervised Learning:** Training a machine learning program by showing examples of input/output pairs in order for the program to map the connection between input and output.

(Kaplan, *Artificial Intelligence: What Everyone Needs to Know*, p. 30)

- **Unsupervised Learning:** Training a machine learning program by giving the program the input data but not initially defining a desired output. This method allows the program to make its own connections throughout the entirety of the input data.

(Kaplan, *Artificial Intelligence: What Everyone Needs to Know*, p. 30)

While these two terms seem to dive unnecessarily deep into the inner-workings of machine learning so early on in the course, it is important to distinguish the capabilities that machine learning programs have and how those capabilities are achieved only through extensive preparation of the program beforehand.

- **Synthetic Intellects:** The product of the efforts of machine learning, big data, neural networks, etc.; systems that recognize patterns from vast amounts of data in order to make real time decisions based on input from an array of sensors monitoring aspects of the world. (Kaplan, *Humans Need Not Apply*, p. 4-5)
- **Forged Laborers:** embodied systems capable of tirelessly performing an astonishing range of chores in chaotic, dynamic environments. (Kaplan, *Humans Need Not Apply*, p. 6)
- **Electronic Entities:** any combination of synthetic intellects and forged laborers, used as a term to directly compare to humans (Kaplan, *Humans Need Not Apply*, p. 6)
- **Automation:** The utilization of synthetic intellects and forged laborers that allows for the accomplishment of more and more complex processes and jobs. (Kaplan, *Humans Need Not Apply*, p. 6)

The final four terms come directly from Kaplan's *Humans Need Not Apply* and aren't necessarily terms that are used throughout the industry, but they are very useful in serving as an encompassing capstone for the products of machine learning. As the course moves forward into the discussion of automation, students will be able to reference the general concepts of "Synthetic Intellects" and "Forged Laborers" instead of having to carefully distinguish the many different methods and uses of machine learning.

Continuing on the topic of terminology, the course will also take into account the origin of some of these terms and the effect these terms have had on the history of AI development. For example, the term "artificial intelligence" was first coined by John McCarthy in a 1955 Dartmouth research proposal, and while the definition has changed throughout the decades, for

better or for worse the name has persisted. In his book *Artificial Intelligence: What Everyone Needs to Know*, Jerry Kaplan explains that the term "artificial intelligence" has significantly influenced our view of the field:

To better understand how the aspirational connection between machine and human intelligence clouds and colors our understanding of this important technology, imagine the confusion and controversy that powered flight might have suffered if airplanes were described from the start as "artificial birds." This nomenclature would invite distracting comparisons between aviation and avians, sparking philosophical debates as to whether airplanes can really be said to "fly" as birds do, or merely simulate flying. (The parallel here is the ongoing debates as to whether machines can really think or just simulate thinking. And the answer is the same: it depends on what you mean.) Yes, airplanes have wings, which were plausibly inspired by bird wings, but they don't flap or fold and the propulsion system is completely different, as is their range, altitude, and just about everything else about them. If this misplaced framing had persisted, there might have been conferences of experts and pundits worrying about what will happen when planes learn to make nests, develop the ability to design and build their own progeny, forage for fuel to feed their young, and so on. As ridiculous as this sounds, its similarity to the current wave of concern about superintelligent machines and runaway AI posing a threat to humanity is stronger than a casual observer might expect. Little or nothing in the field of AI today, other than wild speculation, supports these concerns--at least for the foreseeable future. And if it ever does, we're likely to have plenty of warning.

Had McCarthy chosen a more pedestrian term that didn't suggest a challenge to human dominance or cognition, like "symbolic processing" or "analytical computing," you

might not be reading this book right now. Progress in the field might have merely seemed like what it is--the continuing advance of automation.

(Kaplan, *Artificial Intelligence: What Everyone Needs to Know*, p. 16-17)

The point that Kaplan is arguing is that the wording of this technology really does matter, and so these final five definitions are all very similar in nature but very particular in their wording. All of these definitions have their uses, though, and will be referenced frequently throughout the course when describing the behavior of particular AI programs and scenarios.

- **Artificial Intelligence (1955 definition):** "Making a machine behave in ways that would be called intelligent if a human were so behaving."

(McCarthy, et al. 1955)

- **Artificial Narrow Intelligence:** AI that specializes in *one* area.

(Urban 2015)

- **Artificial General Intelligence:** A computer that is as smart as a human *across the board*—a machine that can perform any intellectual task that a human being can.

(Urban 2015)

- **Artificial Super Intelligence:** An intellect that is much smarter than the best human brains in practically every field, including scientific creativity, general wisdom and social skills.

(Bostrom 1998)

- **Weak AI:** A machine that simulates intelligence

(Kaplan, *Artificial Intelligence: What Everyone Needs to Know*, p. 68)

- **Strong AI:** A machine that *is* intelligent

(Kaplan, *Artificial Intelligence: What Everyone Needs to Know*, p. 68)

The next concern of this week is testing the students on their knowledge of these terms. After going through all of these terms and the relevant information and history involved, students will be given a weekend to study the vocabulary and will then be given a test consisting of two sections: the first section will involve matching the first 13 terms and definitions together from a scrambled list. The second section will give the students the final five terms and require them to write down the definition of each term from memory. The reason here is that the early terms are all relatively simple concepts and vary enough from one another that memorization of the particular wording involved would be unnecessary; the importance is placed on the general understanding of the concepts involved. As for the final five definitions, the specific wording of how we define artificial intelligence has been impactful in the perception of the technology and so it is imperative that students focus more here on the specific wording of these five definitions of artificial intelligence.

Finally, students will also be assigned to read the introduction and first two chapters of *Humans Need Not Apply*, as this book will be the main focus of the class for the coming weeks. The introduction of the book does a great job of explaining why it was written and what it hopes to accomplish by initiating informed discussion on the topic of automation in the workforce. Kaplan explains that "our public discourse lacks the concepts and exemplars to properly describe what's likely to happen as technological progress accelerates, much less to guide us to reasonable solutions" (Kaplan, *Humans Need Not Apply*, p. 15). Throughout the rest of the introduction, Kaplan briefly explains the technologies involved in the development of advanced automation lays out the many reasons why we should be worried about the progress that will soon be made in automation. Some of his warnings are that the current trend of the wealthiest 1 percent of people may be even further reduced to the 0 percent (Kaplan, *Humans Need Not Apply*, p. 11),

that relying on these intricate systems to maintain our daily way of life makes us vulnerable to "havoc on an unimaginable scale in the blink of an eye" (Kaplan, *Humans Need Not Apply*, p. 7), and that our current laws have been constructed with the idea that people can make reasonable decisions like disregarding a small rule in order to save a life, but this will be insufficient as autonomous systems become more widely implemented (Kaplan, *Humans Need Not Apply*, p. 10). Ultimately, Kaplan impresses on the reader that "unless you understand what's really happening, you can't appreciate what's likely to happen in the future" (Kaplan, *Humans Need Not Apply*, p. 13), and that is exactly what students should keep in mind throughout this course, to understand the technology and influence of AI in order to be more prepared for their own future.

The first two chapters of *Humans Need Not Apply*, titled "Teaching Computers to Fish" and "Teaching Computers to Heel" delves into the history of artificial intelligence research and development. Kaplan describes the initial endeavors of the ambitious mathematicians and computer scientists of the mid-1960's as they established the new field. He also explains how it was only a few years later that the higher-ups of IBM were having to quell the fears of managers by making the statement that would limit AI research for decades to come: "computers can do only what they are programmed to do" (Kaplan, *Humans Need Not Apply*, p. 20). While this statement may have been true for computers at the time, researchers and programmers would soon come to find out that the systematic approach of "if this, then that" is not the only way to program a computer (Kaplan, *Humans Need Not Apply*, p. 25). The other factor of the history of AI that Kaplan chooses to cover is the exponential development in technologies involving data management and processing speeds, and he explains how impactful exponential development really is by comparing it to the linear development of gas mileage improvements in cars. It was because of these two main developments that automation was able to leap off of the path of the

relatively simple industrial automation. Ultimately, these chapters clearly review the many shifts in AI technology to explain how and why things are the way they are today, which is essential for this course as the discussion shifts to future impacts and developments.

Week 3

Week #	Topic	Reading	Assignments and Essays	Weekly Objective
Week 3	Automation <ul style="list-style-type: none"> Economics 	<p><i>Humans Need Not Apply</i>: Chapters 3 & 4 (p. 51-74)</p> <p>Video: CGP Grey's <i>Humans Need Not Apply</i> YouTube Video (in class)</p>		Students should be able to understand the rapid impact that narrow AI will have on our unprepared job markets, taking special note as to how it pertains to them individually as they transition into their own careers. They should also be able to examine potential solutions to the expected turmoil in that this accelerated job displacement will cause in the equities of society. This will transition smoothly into the current issues and potential solutions involving education reform.

The first major aspect that will be covered in the effects of widespread automation will be about the impact it will have on our economics as a whole. The major theme to keep in mind throughout this week will be job market disruption due to the competition between humans and electronic entities. Chapters 3 and 4 of *Humans Need Not Apply* will be an introduction to what happens when humans have to compete with electronic entities for the same resources, and how it's not really a competition at all.

The first example Kaplan uses to illustrate this disadvantage is through something called "high-frequency trading (HFT)" on the stock market, which Kaplan explains is the buying and selling of stocks at a rate of 100,000 transactions in a fraction of a second (Kaplan, *Humans Need Not Apply*, p. 52). And while Kaplan goes on to explain how these algorithms operate and why they are having such an impact on our financial markets, he also simplifies the situation by likening HFT's to an unobtrusive robot that walks behind people in a town and picks up any pennies they happen to drop. When an entrepreneur tries to pitch this idea to the town's mayor,

all he really has to point out is that the robot will be cleaning up the town free of charge and that no one will really miss the pennies that they drop. And this is how HFT's really work: they slink around the stock market looking for tiny net gains they can easily make, which are trades that no one else would bother wasting their time to capitalize on. Something to note, though, is that while the streets of the town might be cleaner, the entrepreneur never specified where the collected money was going and the mayor never asked, and as such the entrepreneur becomes much richer. Kaplan offers several solutions to the real world HFT problem, and students should be able to use these solutions to discuss in class some of the potential solutions to the similar penny-pinching robot scenario.

The HFT example serves as great illustrator of two things: 1) people cannot and should not have to compete with electronic entities for the same resources, because it's an unfair fight, and 2) the owner of the electronic entity in question has a huge advantage over people that don't have access to the same technology. The second point is a brief introduction to the topic of social equity and widening wealth gaps in our society, which will be covered later in week 6. For now, after being shown an example of imbalanced competition between humans and electronic entities, students will move forward to chapter 8 to read about how "47% of total U.S. employment is at high risk of significant automation," (Kaplan, *Humans Need Not Apply*, p.152).

Finally, students will be shown a video from the YouTube channel CGP Grey that was posted directly in relation to Kaplan's book of the same title. While Kaplan's full book covers a fairly wide range of topics, Grey only summarizes the job displacement concerns brought to light by Kaplan when he points out that nearly half of our current work force is at risk of automation. Grey does a great job of making the viewer aware of the facts of the situation, but he doesn't go into some of the suggested solutions that Kaplan discusses in his book. Therefore, this video

serves as a great discussion prompt for what students might come up with as potential solutions without their being made aware of potential answers that will be covered next week in chapter 8.

Week 4

Week #	Topic	Reading	Assignments and Essays	Weekly Objective
Week 4	Automation cont. <ul style="list-style-type: none"> ● Education 	<i>Humans Need Not Apply</i> : Chapter 8 (p. 131-158)	Assignment #3: Find something about your career where automation might threaten your ability to get a job. <i>BlackBoard submission</i>	Students should be able to recognize the costs and benefits of the current education system and how the rapid development of automation will influence any changes made to our education system. They should also be able to analyze why potential solutions might help improve the situation when it comes to working with automation instead of against it.

The main reason for skipping forward to chapter 8 of *Humans Need Not Apply* is to establish investment for each student as to how automation will affect them personally. The assignment for this week is also designed to motivate the individual student as well; students will be expected to research ways in which automation advancements might affect their ability to get a job in their desired profession. While similar to the first week's assignment in that students will be more aware of what AI and automation has to offer, the goal of this week's assignment is to push students towards concern about what's to come rather than simply stimulating interest. It's easy to recognize that jobs such as farming and transportation will be the first to be overwhelmingly automated, but people rarely think about how their own jobs might be automated as well.

Following the discussion of the economic impacts of automation, this week will shift the topic slightly towards the structure of our current education system. Students will investigate why the current education system is designed as it is and how it currently attempts to prepare people for the modern job market. Kaplan points out the two main mistakes being made in our current education system: 1) we are entrusting traditional schools with deciding what to teach our

students, and they do not update their curriculum often enough to meet market demand, and 2) we have locked ourselves into a system that front-loads children and young adults with years of schooling, expecting them to get a job, and never to go to school again (Kaplan, *Humans Need Not Apply*, p. 152-153). The point of reading this chapter is to give students an opportunity to think about how we might change our current education system to better prepare people for the rapid job displacement that automation is beginning to cause.

Kaplan proposes that we treat education like we do mortgages. With a house mortgage, a homeowner gets to live in the house they are paying for with the expectation that if they can't make the payments then the property gets returned to the bank (Kaplan, *Humans Need Not Apply*, p. 153-154). A "job mortgage" would work similarly in that the cost of receiving training for a specific job would be paid out initially by the employer (if they aren't providing the training themselves) while the employee would pay for the training directly from a percentage of their salary. This idea differs from the current pretense of relying on for-profit colleges to teach the required skills, which is something that rarely happens and results in people drowning in student debts (Kaplan, *Humans Need Not Apply*, p. 154-155). Kaplan explains that "by creating the proper economic incentives for employers, lenders, and trainers through appropriate public policies, we can render the process of skill acquisition and retraining both practical and humane, not to mention much more effective than it is today," (Kaplan, *Humans Need Not Apply*, p. 156). The education system that Kaplan proposes will give students a better idea of how the job displacement problem can be solved directly through education reform rather than controversial, widespread proposals like universal basic income.

Week 5

Week #	Topic	Reading	Assignments and Essays	Weekly Objective
--------	-------	---------	------------------------	------------------

Week 5	Automation cont. ● Law	<i>Humans Need Not Apply</i> : Chapter 5 (p. 79-92)	Essay #1: Solve a hypothetical court case involving AI/Automation.	First, students should be able to recognize the impacts that automation will have on how the world of law operates, including how lawyers will come utilize it as well as how it may come to be used in the court itself. Then, students should be able to analyze a hypothetical court of law situation involving an electronic entity and be able to come to a "verdict" themselves using at least one real court case as precedence. Following this, students should be able to recognize the importance that legal precedence will have on such a powerful technology.
--------	---------------------------	---	--	--

The fifth week of the course serves as glimpse into hypothetical discussions involving artificial intelligence, something that will be utilized more heavily in the second half of the semester. While the hypothetical nature of this week stands apart from the weeks that surround it, the topic of automation's potential to impact highly professional jobs can be illustrated through the example of automation and law. As mentioned in CGP Grey's *Humans Need Not Apply* video shown in week 3, the work that lawyers do can mostly be automated, and a book later on in the semester will examine the potential for AI to operate as an unbiased judge. Therefore, class discussion can and should flow easily back into chapter 5 of Kaplan's book wherein he discusses electronic entities in the courtroom, not as a judge or a lawyer, but as a defendant.

Kaplan introduces the chapter by explaining the concept of "moral agency" in that moral agents must be capable of "perceiv[ing] the morally pertinent consequences of their actions, and... be able to choose between the relevant courses of action," (Kaplan, *Humans Need Not Apply*, p. 79). While we might immediately assume this describes human behavior, it's important to note that this definition doesn't care how an agent internalizes a situation when it needs to make a decision, only that it is able to react in accordance to its knowledge of the consequences that would follow its decisions. In fact, Kaplan gives two examples of non-human entities being charged with crimes because we believed them to display the necessary qualities of being a moral agent: animals and corporations. In the Middle ages, there are accounts of animals being

tried for criminal offenses, which seemed normal at the time because people believed animals were able to tell right from wrong (Kaplan, *Humans Need Not Apply*, p. 79). We have since come to the understanding that animals should not be tried as moral agents; modern legal theory, however, gives us a more contemporary example of non-human moral agents in the form of corporations. Kaplan explains that when 11 workers died in a BP oil rig explosion in 2010 the US government filed both criminal and civil charges against the BP Corporation. It was argued that "[t]he corporation itself, as distinct from its employees, had a duty to put sufficient controls in place to avoid accidents like this one," (Kaplan, *Humans Need Not Apply*, p. 80). The result is that certain non-human entities can absolutely be labeled as moral agents, and as such can be held responsible in the event that a crime is committed.

As soon as this rhetoric is established for students, the question then becomes whether or not synthetic intellects should be considered moral agents. Kaplan uses the rest of the chapter to discuss this notion, outlining a useful example of a situation involving a synthetic intellect being charged for a crime. The importance of this example lies not just in the hypothetical events of the crime itself, but in how a judge might actually come to a decision based on actual legal precedent. The essay assigned this week will be the first of the three major essays that each student will write during the semester. For this assignment, students will be given another hypothetical situation involving a synthetic intellect "committing" a crime, similar to what Kaplan lays out. Students will be expected to argue why the synthetic intellect qualifies as a moral agent, propose what a judge might decide based on actual legal precedent, decide on a proper sentence for the synthetic intellect and any other parties involved, and explain the effects that this decision could have on future development of similar technology. The goal is to get

students to realize how much our legal system is dependent on legal precedent and how our societal thought, values, and beliefs can (and should) affect technological progress.

Week 6

Week #	Topic	Reading	Assignments and Essays	Weekly Objective
Week 6	Automation cont. <ul style="list-style-type: none"> ● Social Equity ● Solutions 	<i>Humans Need Not Apply</i> : Chapters 7, 9, & "Outroduction" (p. 109-127, 161-208)		Students should be able to recognize how the current automation technologies are impacting social equity by looking at the amount a large corporation makes per employee and comparing it to that company's average employee salary/wage. Students should then be able to deduce the reasons behind particular ideas from proposed solutions to the equity crisis that automation could bring on.

This week sets to finish out Kaplan's *Humans Need Not Apply* by having students read about the ever widening wage gap before they read the final chapter, which is dedicated to some proposed solutions to all of the problems discussed thus far. Chapter 7, titled "America, Home of the Brave Pharaohs," mainly covers the topic of the wage gap by giving two examples of livelihood in America, one at the high end of society and one at the lower end. First, Kaplan elaborates on his own personal lifestyle and what he and his family are able to afford at this point in their lives. In short, he describes himself as being pretty well off, owning an extravagant property and a house that could easily hold 150 people for a party, and for bigger crowds they also have a guest house that can accommodate another 200 people (Kaplan, *Humans Need Not Apply*, p. 111). As a side note, Kaplan explains that it hasn't always been like this and that they have had to work their way up from warehouse jobs and subfreezing temperature, roach-infested, studio rentals in Brooklyn. Even so, Kaplan is sure to make clear to the reader that their current lot in life doesn't even bring them close to the wealthiest top 1% of Americans (Kaplan, *Humans Need Not Apply*, p. 111). He goes on to illustrate what the wealthiest are capable of accomplishing with their money, and that Jeff Bezos's current assets allow him to "[make] more

on a Saturday spent on the golf course than the other college grads in his foursome, taken together, will earn in their entire lifetimes," (Kaplan, *Humans Need Not Apply*, p. 112). In contrast to Kaplan's own story, he then shares the story of Emmie Nastor, "the perfect employee," whom he hired in 2009 as a receptionist for their small game company operation. Emmie's entire story is an endearing example of how perseverance, diligence, and hard work doesn't always lead to a better life.

Emmie's parents were immigrants who worked extremely hard to put him through 8 years of part-time college in order to give him access to what they hoped would be a better life than they had endured. His father worked initially as a U.S. Air Force mechanic in the Philippines and after emigrating to the States was employed by a telecommunications company installing landlines, and as a father he did everything he could to ensure that his children would be able to do more. Following his graduation from San Francisco State University, Emmie spent over 8 hours applying to 20-30 job opportunities every single day for three months straight, amounting to over 1,800 applications before finally landing his first interview – Kaplan's office receptionist position. Emmie was an excellent and hardworking employee, always demonstrating that no task was beneath him and never voluntarily late or absent to work. Emmie explained that the work was enjoyable and the job served him well until they were bought out by another larger company and his job was deemed irrelevant. Fast forward through another couple of months of applications and job searching and Emmie finally landed another job... at the same telecommunications company that employed his dad, doing the same installation that his dad has spent his life doing. The work is grueling and the hours are extensive, but Emmie didn't complain and put forward as much effort as could, until the point that he literally breaks his back while performing an installation, causing him to be pushed to an office job that could easily be

replaced by automation (Kaplan, *Humans Need Not Apply*, p. 124-125). The result of this story is that Emmie's household income, which includes his wife and brother's contributions, amounts to \$53,046, which actually falls above the national median, leaving him better off than over half the people in the nation even though it doesn't feel like it (Kaplan, *Humans Need Not Apply*, p. 126).

These two examples appear as stark contrasts when put side by side, but the main point that Kaplan tries to make to the reader is that these aren't even on opposite ends of the income spectrum in America, and that the situation is much worse for over half of the people in America. This chapter stands out from the others because the wage gap in America isn't something that can be blamed on automation or AI, and Kaplan doesn't attempt to elaborate on the history of the subject either; instead, he emphasizes that the situation is real and urgent. Recognition of the situation at hand will be important as students move forward in the semester, keeping in mind how automation and AI advancements may only worsen the current climate. The final chapter attempts to calm the storm a bit by acknowledging the reality of the problems in question, such as the fact that automation technologies will continue to advance, that people will be heavily displaced in the workforce by automation, and that our current education system is not prepared for the constant retraining necessary to keep the average American worker relevant. If nothing is done to prepare our society for these imminent changes, then today's current wage gap will only grow exponentially, damning our socioeconomic climate to years of hardship before it is able to rebalance, if it can. It is important for students to keep in mind that Kaplan's solutions aren't the only improvements that can be made; rather, Kaplan is simply proposing a collection of ideas to go along with the warnings that he has spent the entire book bringing to light. The purpose of *Humans Need Not Apply* is to bring about conversation on the topic of automation and what all it will entail for every American citizen.

Week 7

Week #	Topic	Reading	Assignments and Essays	Weekly Objective
Week 7	Big Data and Machine Learning <ul style="list-style-type: none"> Intro to schools of thought on AI development 	<i>The Master Algorithm</i> : Prologue		Students should be able to recall the major schools of thought on how AI is being approached from a developmental aspect. Students should then be able to recognize what school of thought is being utilized in a given example of research and development on artificial intelligence.

The entire course is separated into two major parts: the first half addresses the economics and social equity of what's to come with advancements in artificial narrow intelligence, while the second half will look into what artificial general intelligence will bring about as we must potentially redefine what it means to be human. However, before those broad topics can be discussed, it is important to inquire about research and development on the transition from ANI to AGI, and why the transition has been and will continue to be a difficult step in the road to AGI. These questions are addressed well enough by Pedro Domingos in the prologue of *The Master Algorithm*, which first briefly describes the widespread use of machine learning algorithms and explains the great advantage of this relatively new tool to humanity. Domingos clearly introduces machine learning to people who may never have heard of it, and he claims that it is such a powerful and pervasive tool that "it's not wise to let it remain a black box... [o]pacity opens the door to error and misuse... You can't control what you don't understand, and that's why you need to understand machine learning--as a citizen, a professional, and a human being engaged in the pursuit of happiness," (Domingos, *The Master Algorithm*, p. xvi). He then goes on to clarify that there are five main schools of thought when it comes to machine learning development. Here is how Domingos explains each:

Symbolists view learning as the inverse of deduction and take ideas from philosophy, psychology, and logic. Connectionists reverse engineer the brain and are inspired by

neuroscience and physics. Evolutionaries simulate evolution on the computer and draw on genetics and evolutionary biology. Bayesians believe learning is a form of probabilistic inference and have their roots in statistics. Analogizers learn by extrapolating from similarity judgements and are influenced by psychology and mathematical optimization.

(Domingos, *The Master Algorithm*, p. xvii)

All five of these schools of thought have their own theoretical general-purpose learner which Domingos terms as a "master algorithm," in which the algorithm could "discover knowledge from data in any domain," (Domingos, *The Master Algorithm*, p. xvii). Each one of these "master algorithms" have their advantages and disadvantages and Domingos believes that only a special combination of each can result in what he calls "the ultimate master algorithm" (Domingos, *The Master Algorithm*, p. xvii). Domingos claims that "[i]f it exists, the Master Algorithm can derive all knowledge in the world--past, present, and future--from data... The Master Algorithm is to machine learning what the Standard Model is to particle physics or the Central Dogma to molecular biology: a unified theory that makes sense of everything we know to date, and lays the foundation for decades or centuries of future progress," (Domingos, *The Master Algorithm*, p. xviii). Domingos delves farther into each school of thought later in the book, however, the introduction introduces the topic of machine learning well enough that reading any more of the book would be unnecessary for the range of this course. The rest of the week's instruction will delve into the details and operations of certain machine learning techniques in order to explore the black box issue for the students.

Week 8

Week #	Topic	Reading	Assignments and Essays	Weekly Objective
--------	-------	---------	------------------------	------------------

Week 8	Review of part I and introduction to part II			
--------	--	--	--	--

Week 8 marks the halfway point of the semester and will be utilized as a review/recap period for the previous 7 weeks and as an introductory period for the topics to follow in the remaining weeks. This week can also act as leeway in the schedule should some of the previous topics take longer than expected to cover; equally, should the course be running ahead of schedule, week 9's material can be pushed forward in order to stretch out some of the more dense topics to come. It is important to note that from this point forward discussion will be based on consciousness forming within AI. Therefore, unless otherwise specified, "AI" will refer to artificial general intelligence (AGI) or artificial super intelligence (ASI) as it is agreed that consciousness can not form within artificial narrow intelligence.

Weeks 9 & 10

Week #	Topic	Reading	Assignments and Essays	Weekly Objective
Week 9	What is Consciousness? <ul style="list-style-type: none"> History of Philosophy of Consciousness Can it be programmed? What happens then? 	<i>In Our Own Image: Savior or Destroyer?</i> Introduction & chapters 1-3 (p. vii-xxi, 3-47)	Assignment #4: Prior to starting this half of the course, mini essay on what the student thinks it means to be conscious, intelligent, and alive. <i>BlackBoard submission</i>	Students should be able to recognize the role of consciousness in the past and future of artificial intelligence. They should also be able to recognize how the philosophy of consciousness and its history has influenced the perception of AI in comparison to ourselves. Students should also be able to recognize the difficulties faced in trying to program consciousness (including the difficulties faced in testing for its existence in other beings) as well as what the possible ramifications we might see should it ever be accomplished.
Week 10	Consciousness and AI <ul style="list-style-type: none"> History of anthropomorphism History of how we define the mind 	<i>In Our Own Image: Savior or Destroyer?</i>	Assignment #5: Mini essay; choose one of the five major paradigm shifts covered in class and come up with a metaphor of your own for the mind-body problem based on the technology of the time period in question. (Note: "technology" is not restricted to "inventions," rather it refers to the science and knowledge utilized for society's advancements) <i>BlackBoard submission</i>	Students should be able to relate our history of consciousness-philosophy to our long history of anthropomorphism in relation to contemporary technology. Students should also be able to recall the major movements throughout history in how we have defined our mind as well as how the technology of the time heavily influence our definition of ourselves.

Starting in weeks 9 and 10, students will begin to look at artificial general intelligence in greater detail and with the added perspective of philosophy, differing from the stricter grounded viewpoints of the previous 8 weeks. George Zarkadakis argues for the necessity of philosophy in AI development as well as illustrating the connections between our history and our future when it comes to technology and philosophy of the mind and body. Students should recall Kaplan's argument for how the particular naming of "artificial intelligence" creates a notion to anthropomorphize inanimate objects when they begin to exhibit behaviors similar to our own. Zarkadakis explains that this phenomenon arose in the modern evolution of our species:

The advantages of anthropomorphising when hunting become clearly, if not dauntingly, apparent. Modern humans, by imagining animals as possessing thoughts, could predict animal behaviour better... Their modern way of thinking, equipped with an advanced theory of mind, reaped clear utilitarian benefits from anthropomorphising animals... The connection between survival and imagining non-human minds was forged forevermore. (Zarkadakis, *In Our Own Image*, p. 22)

Most of part 1 of Zarkadakis' book references our anthropomorphising habits in relation to what is often referred to in the field of philosophy as the mind-body problem. Zarkadakis explains that throughout history, humans have tried to understand our brain and our individual consciousness by means of metaphors relating the body and the brain to concepts that are easier to understand and relate to (Zarkadakis, *In Our Own Image*, p. 29). Something to note is that each age has its own metaphor for the brain and that this metaphor can change and adapt along with society; this change is known as a "paradigm shift" (Zarkadakis, *In Our Own Image*, p. 29). Furthermore, we have seen around five to six major paradigm shifts in the history of the mind-body problem, each of which has been driven by the dominant science and technology of the

time (Zarkadakis, *In Our Own Image*, p. 29). The first defining metaphor was likening humans to mud, which reflects the adoption of agriculture and the multitude of creation myths involving a deity molding the first humans from mud and breathing life into them (Zarkadakis, *In Our Own Image*, p. 29-30). The first paradigm shift of the mind body problem didn't arise until the Greeks invented hydraulic systems that allowed for crude movement of inanimate objects. Our susceptibility to anthropomorphising these simple machines led us to believe that it was the fluid within ourselves that gave us life, and that introducing fluid to other objects would imbue them with similar characteristics (Zarkadakis, *In Our Own Image*, p. 31). Zarkadakis goes on to list the other couple of paradigm shifts, mechanical and bioelectrical, before reaching our modern-day metaphor in which we liken the brain to a computer and vice-versa. It is this defining metaphor along with our predetermination to anthropomorphize our technologies that has driven much of the concern and interest within AI development.

The first three chapters of Zarkadakis' book do not bring up any difficult concepts; instead, they help the reader recognize the difficulty in defining the brain and how easy it is to turn to our current technology in order to better understand ourselves. Our usage of metaphors isn't unique to the conceptualization of the brain. In fact, Kaplan's definition of "learning" from week two actually refers to repeatedly using metaphors in order to better understand the world. This concept will become more important as students begin to look for examples of machines learning through metaphors in order to display their individual consciousness.

The assignment for week 9 is for students to write one to two pages on what they think it means to be alive, being sure to explain where and if consciousness fits into their ideas. Students will complete this assignment prior to the reading in order to help them see how common it is to relate themselves to modern technology before being told why we tend to do so. This assignment

could also be altered slightly to require students to follow up on their original understandings after reading on consciousness. The assignment for week 10 will be for students to choose one of the five major paradigm shifts in the mind-body problem and to come up with a different metaphor that references another one of that time period's contemporary technology. Students should explain the metaphor in full, making direct connections between the defining characteristics of the technology and the relevant characteristics of our mind and body. Students should also keep in mind that "technology" is not restricted to "inventions," rather it refers to contemporary science and knowledge utilized for society's advancements; it is the widespread use of the technology that allows society to better understand what it means to be human by likening a thing they understand to a concept that is hard to grasp.

Week 11

Week #	Topic	Reading	Assignments and Essays	Weekly Objective
Week 11	Consciousness and AI <ul style="list-style-type: none"> ● Responsibilities as creators ● Programming ethics into a machine 	<i>Frankenstein</i> by Mary Shelley	Assignment #6: Mini essay on personal viewpoint in regards to what Frankenstein has to tell us about the responsibility we have when it comes to developing powerful technology. <i>BlackBoard submission</i>	Students should be able to recognize the current stances on what precautions and safeguards should be taken when developing AI. Students should be able to make an argument regarding what our responsibilities as creators should be if we continue to pursue the development of artificial consciousness. This argument can and should utilize a blend of contemporary technology and arguments as well as pieces of science fiction that portray the dangers we pose to ourselves when technology gets out of our hands

Mary Shelley's *Frankenstein* was influenced by Luigi Galvani's experiments on bioelectricity, a technological advancement that caused one of the five major paradigm shifts. The tale is told from the perspective of Victor Frankenstein, a scientist who has uncovered the secret to creating life. While Shelley doesn't fully elaborate on the specifics of the experiments that Frankenstein performs, the result remains that an ambitious scientist creates life in the form

of what he calls a monster. The rest of the story provokes many questions as Frankenstein doesn't know how to best handle the conflicting responsibilities he has to his creation and to the rest of humanity when he realizes that the monster has the capability to bring harm to the world.

For assignment #6, students should draw direct comparisons to similar questions involving the emergence of AGI and ASI such as, "how much of a guiding hand should we have in how AI is maintained should it become a self-conscious entity?" or "should we always pull the plug when a program starts to show signs of self-consciousness because of the potential dangers to humanity?" Students should be expected to discuss what Mary Shelley has to say about the topic through *Frankenstein* as well as to give their own opinions on the subject based on what they have learned so far about AI.

Another topic to keep in mind is the programming of ethics into an AI. It wasn't possible for Frankenstein to preemptively instill ethics into his creation because it was born tabula rasa; the creation's moral understandings could only come from what it learned about the world through its own experiences. Programming gives us as creators an advantage as it allows us to put in safeguards to control how our creations are able to operate. Students will be shown several of the current endeavors in ethics programming and we will discuss the problems we face in deciding the particulars of the ethics a program should follow.

(Note: prior to starting week 11, students will have been advised to begin reading *Frankenstein* early so that their reading load isn't as heavy during week 11 itself.)

Week 12

Week #	Topic	Reading	Assignments and Essays	Weekly Objective
--------	-------	---------	------------------------	------------------

Week 12	Consciousness and AI <ul style="list-style-type: none"> • What does society look like with another form of intelligence/consciousness? 	<i>I, Robot</i> by Isaac Asimov Stories to read: "Robbie," "Runaround," "Reason," "Liar" (p. 1-81, 111-135) "Evidence" "The Evitable Conflict" (p.206-272)	Essay #2: Comparative essay between I, Robot and Frankenstein	Students should be able analyze the potential impacts on society should another form of intelligence/consciousness begin to operate alongside ourselves. Things that should especially be noted are laws and regulations, societal norms, and the interactions that might take place.
---------	---	---	---	---

Isaac Asimov's *I, Robot* is a collection of short stories, each of which gives a unique perspective on how society performs when a secondary high-level consciousness must coexist with humanity. Because almost every story can be read as in individual piece, students won't be required to read the novel in its entirety. The first story ("Robbie") illustrates the difference in how a child growing up alongside AI might view the creation as opposed to how an older generation that experienced the gradual progression from ANI to AGI. The next three that students should read will be "Runaround," "Reason," and "Liar," which introduce Asimov's three laws of Robotics and experiments with them through several different logical/ethical scenarios. The three "Rules of Robotics" are as follows:

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey orders given it by human beings except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

(Asimov, *I, Robot*)

These laws are a great example of the previously discussed programming of ethics; Asimov writes that "the three Rules of Robotics are the essential guiding principles of a good many of the world's ethical systems," (Asimov, *I, Robot* p. 221). This quote comes from the

second section of stories that students will read: "Evidence," and "The Evitable Conflict." These are the two final stories of Asimov's collection, expanding the scenarios of the three laws to situations that have a greater impact on a society already coexisting with AI. And while these individual questions are a significant step forward, the most important aspect is that they leave the conclusion of the collection quite open-ended. Even the name of the last story reaffirms Asimov's purpose, to illustrate a future brimming with the potential for conflict, but reminding the reader that we write our own future.

After reading and discussing *I, Robot*, students should be able analyze the potential impacts on society should another form of intelligence/consciousness begin to operate alongside ourselves. Things that should especially be noted are laws and regulations, societal norms, and the interactions that might take place.

The essay assignment for this week will require the students to compare and contrast how these two science-fiction authors portray the outcomes of major advancements in technology. Asimov's world is quite similar to Shelley's as they both delve into the dynamic between the creator and the created, but while *Frankenstein* deals with a first-time creation scenario, Asimov pushes forward to a point where these new entities already have defined roles within society. Each perspective on the dynamic is equally important as humanity dives further into AI technology; how we handle the first emergence of AGI will absolutely define how society accepts and integrates another intelligent life form.

Weeks 13 & 14

Week #	Topic	Reading	Assignments and Essays	Weekly Objective
Week 13	Today's viewpoints on AI <ul style="list-style-type: none"> ● Political? Cultural? ● Where do people stand? 	Individually decide on a movie displaying AI as a		Students should be able to recognize contemporary viewpoints on artificial intelligence in society, including political, cultural, economic, and

	<ul style="list-style-type: none"> Where do you stand now? 	main theme. A list of options will be provided.		philosophical. Students should also be able to look back and see how their own viewpoints have changed throughout this course. They should then be able to come to their own conclusions as to if enough is being done to inform the public (as well as each of these individual areas) on the importance of AI.
Week 14	Popular Culture and AI <ul style="list-style-type: none"> Movies/shows, books, games, etc. 		Essay #3: Write a report on how AI was displayed in the movie you chose. What about it was accurate/inaccurate? How did the culture react to that form of AI? Etc.	Students should be able to analyze a particular piece of popular culture centered around artificial intelligence and be able to report on its accuracy in its portrayal of the technology; they should also be able to analyze what the intent of the piece is when it comes to AI.

While iconic, Shelley and Asimov are only a two examples of how science-fiction can generate conversation on the impacts of technological advancements. Also, *Frankenstein* was first published in 1818 (then called *The Modern Prometheus*) and the stories from *I, Robot* were written throughout the 1940's; the themes and the questions they pose may remain important, but they are hardly current in their handling of popular culture. Students will use their current knowledge and experiences of pop culture, supplemented with class discussion, to construct a preliminary idea of how our society currently perceives AI advancements. Along with this, students will be shown specific examples of statements from contemporary figures such as politicians and entrepreneurs on what should be done about AI. Another important consideration is the concerns of the general public, which is often influenced by pop culture, and whether or not the recent rise in the portrayal of AI in media such as movies and television is impacting how the public views AI research. Students will also look at how their own perception of AI has changed throughout this course, comparing what they already knew to what they didn't know or previously misunderstood. Finally, during these two weeks each student will decide on a movie to watch that utilizes AI as a major theme and write a 3-5 page report analyzing the portrayal of AI within the film. Specifically, students should look for how the movie illustrates specific aspects of AI accurately/inaccurately as well as what questions the movie tries to bring forward

about AI. Students should also look at how the society constructed in the film reacts to the technology.

These various approaches will allow students to paint a more complete picture of how society understands AI. Following this, students will be expected to come to their own conclusions about whether enough is being done to educate the general public on the true nature of AI research and to prepare them for the continuing implementation of the technology in our society.

Week 15

Week #	Topic	Reading	Assignments and Essays	Weekly Objective
Week 15	Overview/Review		Final Presentation: Utilizing everything you've learned this semester, give an individualized presentation expanding on how artificial intelligence will affect your career and life as you move past college and into the world.	

Similar to week 8, the final week of class will serve as an overview of the material covered in the second half of the course as well as acting as leeway in the schedule should any of the topics take more time than expected. Also, students should have begun preparing an individualized presentation on how artificial intelligence will impact them personally as they begin to move into careers and life outside of college. Students are expected to utilize everything they have learned and discussed throughout this semester to answer questions such as "In what ways will the technology that is available now/will be available in the future affect you?" and "How has this course prepared you for the implementation of AI in your field?" This assignment is meant to refocus each student back to their own lives and apply the things learned from this

course to preparing themselves for future development and implementation of AI into society. While there will be time for each student to present during the final exam time-slot, this week can also be allotted for student presentations.

Conclusion

The course that I have designed fits the format of a University Honors seminar and fulfills the learning outcomes of UTC's Thought, Values, and Beliefs general education category. The seminar format allows for class discussion, which is necessary for a course that covers ethical, philosophical, or theological ideas in a developing field. The class conversation and debates that arise from asking the right questions will be more beneficial than any amount of reading or lecturing could ever hope to be for the students involved. The course has been carefully structured to provide understanding of the technology being discussed, and then transitions into the broader areas of social concern. Assignments for the class range from reading explanatory pieces on AI, to philosophical studies on what might constitute life, and to literature/film that might deal with the science and creation of life, such as Mary Shelley's *Frankenstein*. Discussions will be centered on questions such as: What constitutes life, and how does this definition relate to the technology currently in development? What might be the best way to direct research endeavors? And how should AI and its researchers be governed?

I chose to design a general education course on the societal impacts of artificial intelligence in part because of my interests in the topic, but mainly because as I learned more about the potential effects AI will have on everyone's life, the more I realized how little the technology was being discussed by the general public. One of the main issues covered in Kaplan's *Humans Need Not Apply* is that widespread job displacement caused by automation will bring into question the effectiveness of our current education system; college educated students

entering the job market are being trained and educated in skills that could quickly be rendered irrelevant because of the expected constant shifts of job requirements. While I am not able to provide my fellow students with the skills they need to succeed in their careers, in designing this course I hope at least to make them aware of situation they will face.

References

n.d. "Algorithm." *Oxford Dictionaries*. Accessed March 15, 2018.

<https://en.oxforddictionaries.com/definition/algorithm>.

Asimov, Isaac. 2008. *I, Robot*. New York: Bantam Books.

n.d. "Big Data." *Oxford Dictionaries*. Accessed March 15, 2018.

https://en.oxforddictionaries.com/definition/big_data.

n.d. "Black Box." *Merriam-Webster*. Accessed March 15, 2018. [https://www.merriam-](https://www.merriam-webster.com/dictionary/black%20box)

[webster.com/dictionary/black%20box](https://www.merriam-webster.com/dictionary/black%20box).

Bostrom, Nick. 1998. *How Long Before Superintelligence?* October 25. Accessed March 15,

2018. <https://nickbostrom.com/superintelligence.html>.

Domingos, Pedro. 2015. *The Master Algorithm: How the Quest for the Ultimate Learning*

Machine Will Remake Our World. New York: Basic Books.

Giōrgos, Zarkadakēs. 2016. *In Our Own Image: Savior or Destroyer? The History and Future of*

Artificial Intelligence. New York: Pegasus Books LLC.

Grey, CGP. 2014. "Humans Need Not Apply." *YouTube*. August 13. Accessed April 5, 2018.

<https://youtu.be/7Pq-S557XQU>.

n.d. "Hardware." *Oxford Dictionaries*. Accessed March 15, 2018.

<https://en.oxforddictionaries.com/definition/hardware>.

Kaplan, Jerry. 2016. *Artificial Intelligence: What Everyone Needs to Know*. New York: America

of Oxford University Press.

—. 2015. *Humans Need Not Apply: A Guide to Wealth in the Age of Artificial Intelligence*. New Haven: Yale University Press.

n.d. "Machine Learning." *Wikipedia*. Accessed March 15, 2018.

https://en.wikipedia.org/wiki/Machine_learning.

McCarthy, J, M L Minsky, N Rochester, and C E Shannon. 1955. "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence." Hanover, August 31.

Shelley, Mary. 1986. *Frankenstein*. New York: Penguin Books.

n.d. "Software." *Oxford Dictionaries*. Accessed March 15, 2018.

<https://en.oxforddictionaries.com/definition/software>.

Urban, Tim. 2015. "The AI Revolution: The Road to Superintelligence." *Wait But Why*. January 22. Accessed March 15, 2018. <https://waitbutwhy.com/2015/01/artificial-intelligence-revolution-1.html>.