Minor Outline: AI & Humanity

General Introduction

Artificial Intelligence (AI) has been an exciting area of research and innovation for the scientific community for the better part of a century now. But the present moment is particularly important because the influence of AI on our everyday lives has grown exponentially: the way we interact with our bodies, our immediate environment, our societies, and the nation-state is heavily mediated by AI. At this crucial point in the history of AI, students at IIT Hyderabad will benefit from thinking beyond the purely technical aspects of AI, and exploring perspectives offered by philosophers, artists, economists, anthropologists, psychologists and political scientists. By bringing together Engineers, Designers, and Social Scientists at IITH, the AI & Humanity Minor in the Liberal Arts seeks to introduce students to the exciting possibilities and challenges offered by socially significant applications of AI.

Minor Team: Aalok Khandekar (LA), Nandini Ramesh Sankar (LA), Anindita Majumdar (LA), Haripriya Narasimhan (LA), Mahati Chittem (LA), Aravind Kumar Rengan (BME), Sumohana Channappayya (EE), G. V. V. Sharma (EE), Vineeth Balasubramanian (CSE), Prasad Onkar (Design)

AI & Humanity Minor will:

- Satisfy LA credit requirements for B. Tech students
- Be open to participation from Second Year students onwards, since it does not have any prerequisites.

Organization:

- As a series of 1 and 2 credit courses that integrate technical and social perspectives and analyses
- Four kinds of courses:
    - 3 credits equivalent of Foundations courses
    - 3 credits worth of Special Topics courses
    - 3 credits worth of courses focused on social, economic, and governance implications of AI
    - 3 credits worth of project work
- Instruction (partly) through a Digital Humanities Lab, set up as a dedicated instance of the PECE platform which is simultaneously an archival, analytical, and publication space.

Semester 1: Foundations

LA 1500: What is AI & Humanity?
This is the introductory course to the LA minor, AI & Humanity. The course will orient students to the scope of the minor through a series of guest lectures and field trips that expose them to various perspectives on the possibilities and challenges that AI technologies imply. Topics covered include, among others, AI in governance, healthcare, innovation, manufacturing, mobility, and the arts.

Attendance and participation in all course meetings and activities is mandatory for all students. Evaluation will be based on course participation and brief weekly reflection essays.

LA 1470: Philosophical Perspectives on Human-Technology Interactions (1 credit)
This course will look at basic philosophical questions surrounding AI and Humanity:
What does it mean to be 'human' in the 21st century, and how does the growth of AI technologies modify the terms of this question? Can traditional Cartesian dichotomies still be operative in the face of AI? What alternative models of consciousness and cognition do we have to generate in order to address the ontological challenges posed by AI? How does one place the AI debate within larger discussions about the interactions between technologies and societies? Insights from both the Analytic and the Continental traditions in philosophy will be used to address these questions.

References:
- Haraway, Donna. "A Cyborg Manifesto." 1985. http://faculty.georgetown.edu/irvinem/theory/Haraway-CyborgManifesto.html
- Boden, Margaret A., ed. *The Philosophy of Artificial Intelligence*. Oxford University Press, 1990.
- Copeland, Jack. *Artificial Intelligence: A Philosophical Introduction*. Blackwell, 1993.
- Muller, Vincent, ed. *Philosophy and Theory of Artificial Intelligence*. Springer, 2013

LA 1480: Artificial Intelligence in Literature and Popular Culture (1 credit)
The art of the twentieth century and after has moved uncannily close to the impersonality and automation of thinking machines, embracing forms of mediated, augmented, and suspended consciousness provided by algorithms. The central premise of the course is that as a set of practices founded on principles of artifice as well as the idea of a distinctly 'human' creativity, the Arts have a powerful but conflicted affinity with the premises of AI. Using examples from the visual arts, film, and literature, this course will explore the ways in which popular culture has represented Artificial Intelligence as a key domain for reflecting upon the category of the human. The course will also provide students with tools to interact productively with the visual arts and literary/philosophical texts.

References:
- Shelly, Mary. *Frankenstein*. 1824 edition. Project Gutenberg etext. (Extracts)
- Asimov, Isaac. *I, Robot*. Gnome Press, 1950.

- Dick, Philip K. *Do Androids Dream of Electric Sheep*? Doubleday, 1968.
- Rucker, Rudy. *Software*. Ace Books, 1982.
- Baudrillard. *Simulacra and Simulation*. Translated by Sheila Glaser. University of Michigan Press, 1994. (Extracts)
- The Wachowski Brothers. *The Matrix*. Warner Brothers, 1999.
- Tyldum, Morten. *The Imitation Game*. Black Bear Pictures, 2014
- Garland, Alex. *Ex Machina*. Universal Pictures International, 2015.

LA 1490: Socio-Cultural Perspectives on Artificial Intelligence (1 credit)

AI technologies shape not only what humans (can) do as part of their daily routines, but simultaneously also transform their very ideas of what it means to be human and how they relate to each other. In other words, core categories of social life--including for example, self and community, agency and autonomy--are challenged and rearticulated as they come to be increasingly mediated by AI technologies. And yet, as empirical social researchers have convincingly demonstrated, such newer conceptions of self, identity, and community do not arise in a vacuum -- but rather always build on culturally and historically available templates of thought and action. Building on scholarship in the history and social anthropology of science and technology, this course seeks to understand this mutual shaping of AI technologies and humanity.

References:

- Downey, Gary. 1998. *The Machine in Me: An Anthropologist Sits Among Computer Engineers*. New York, London: Routledge.
- Boellstroff, Tom. 2015. *Coming of Age in Second Life: An Anthropologist Explores the Virtually Human*. Princeton, NJ: Princeton University Press.
- Suchman, Lucy. 1987. *Plans and Situated Actions: The Problem of Human-Machine Communication*. Cambridge: Cambridge University Press.

Semester 2: Special Topics

LA XXXX: Artificial Intelligence and Behavioural and Mental Health

How are you today? How do you make sense of this information? What kind of people should I target for my product? These may be questions that will soon be answered by artificial intelligence. Advances in technology are bringing physical and mental healthcare closer to individuals who need it. This course will summarize the need for and advances in AI and behavioural and mental health practice, its significant role in medical decision-making and communication, and how it helps shape targeted service delivery through facilitating segmentation.

Some topics which will be covered through this course are:

1) *AI's contribution in redesigning healthcare* which will deal with topics ranging from medical and health decision-making (e.g., booking medical appointments, role in medical assessments, behavioural change and sustenance) to medical interventions and communication (e.g., telehealth, chatbots, empathy).

2) *AI and mental health* which will discuss how technology (especially bots) address depression, stress and other emotional issues.
3) *Psychographics* which will introduce and describe how AI aids in identifying and categorizing behaviour in order to develop tailored service delivery.

References:

- Luxton, D. D. (Ed.). (2015). *Artificial intelligence in behavioral and mental health care*. Academic Press.
- Ebert, D. D., Van Daele, T., Nordgreen, T., Karekla, M., Compare, A., Zarbo, C., ... & Kaehlke, F. (2018). Internet-and Mobile-Based Psychological Interventions: Applications, Efficacy, and Potential for Improving Mental Health. *European Psychologist*.
- Srividya, M., Mohanavalli, S., & Bhalaji, N. (2018). Behavioral Modeling for Mental Health using Machine Learning Algorithms. *Journal of medical systems*, *42*(5), 88.
- Janssen, M., Scheerder, J., Thibaut, E., Brombacher, A., & Vos, S. (2017). Who uses running apps and sports watches? Determinants and consumer profiles of event runners' usage of running-related smartphone applications and sports watches. *PloS one*, *12*(7), e0181167.
- D'Alfonso, S., Santesteban-Echarri, O., Rice, S., Wadley, G., Lederman, R., Miles, C., ... & Alvarez-Jimenez, M. (2017). Artificial intelligence-assisted online social therapy for youth mental health. *Frontiers in psychology*, *8*, 796.
- Cameron, G., Cameron, D., Megaw, G., Bond, R., Mulvenna, M., O'Neill, S., ... & McTear, M. (2017, July). Towards a chatbot for digital counselling. In *Proceedings of the 31st British Computer Society Human Computer Interaction Conference* (p. 24). BCS Learning & Development Ltd..

LA XXXX and **LA XXXX** that we are still working to finalize. The focus will be on broad domains where AI is projected to have transformative impacts: e.g. Health Care, Education, Transportation, and Manufacturing. Faculty from engineering departments along with external guest speakers will be enrolled for guest lectures, to introduce students to a broad variety of ways in which people are imagining AI technologies and the future of human societies. We will report further developments to the Senate in due time and seek approval for these courses as we finalize them.

Semester 3: Social, Economic, and Regulatory Implications of AI

LA XXXX: AI and Social Justice (1 Credit)

Is AI likely to ameliorate or exacerbate existing social and economic inequalities? What does automation herald for the vast section of Indian laborers that work in the informal economy? In what ways are AI technologies gendered? What new forms of empowerment and marginalization is AI likely to generate? How can we understand implicit biases in how AI technologies are being designed and disseminated? How might we address these issues?

References:

- Eubanks, Virginia. 2018. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin's Press.

- Data & Society, especially on algorithmic accountability

LA XXXX: The AI Economy (1 Credit)
What are the economic implications of large-scale automation? How have such far-reaching economic transitions been addressed historically? What solutions are being proposed to modulate disruptions that AI might lead to? This course will look at similar discussions arising during the period of the Industrial Revolution the introduction of Assembly Line manufacturing as a way to understand the kinds of anxieties that automation has historically provoked, the various models that have been implemented to manage such disruptions, as well as ideas such as Universal Basic Income and Taxing Robots that have been recently been proposed toward mitigating the effects of AI-led automation.

References:

- Ford, Martin. 2016. *The Rise of the Robots: Technology and the Threat of Mass Unemployment*. New York: Basic Books.

LA XXXX: Governance Perspectives on AI (1 Credit)
What kinds of governance challenges does AI present? How do ideas such as (human) agency and liability change in the face of increased automation? Are existing legal, regulatory, and governance frameworks sufficient to address these challenges? What ideational, institutional, and practical innovations of jurisprudence will be needed in order to take on the many challenges that AI technologies potentially present?

Semester 4: Project

LA XXXX Project (3 credits): Projects that groups of students will undertake under joint supervision of participating faculty. Projects will be carried out on a Digital Humanities platform that we will set up specially for this minor. The platform will as an archival, research, and publication space for the proposed minor in AI & Humanity.