Ethical Implications of AI for Entertainment and Education – Communication and Philosophical Perspectives

Thursday April 10, 2025 | BU Hillel Rm. 426 | 213 Bay State Rd.

Friday, April 11, 2025 | Wightman Mansion, Hewitt Boardroom | 43 Hawes St.

Sponsors: BU Center for the Humanities | Center for Data Sciences | College of Communication | Division of Emerging Media Studies | Feld Professorship of Emerging Media Studies

BOSTON UNIVERSITY

Table of **Contents**

Welcome to **Beyond the Binge**

From the organizers

The pervasive influence of algorithms on our entertainment and educational consumption is undeniable, serving at times to increasingly blur and draw novel lines between them. Thank you for joining us in grappling with complex questions regarding futures of education and entertainment that will inevitably be shaped by AI. We are so grateful for your commitment to fostering a critical dialogue among computer scientists, media theorists, philosophers, humanists and historians with the aim of generating actionable recommendations for a more ethical and equitable future.

James E. Katz, Feld Professor of Emerging Media Studies
Juliet Floyd, Borden Parker Bowne Professor of Philosophy; Director, Boston University Center for the Humanities (BUCH)

Co-Sponsors

- BU Center for the Humanities
- BU Center for Data Sciences
- BU College of Communication
  - Feld Professorship of Emerging Media Studies
  - Division of Emerging Media Studies

The workshop organizers express particular gratitude to COM Dean Mariette DiChristina-Gerosa for her generous support.

International
Steering Committee

James E. Katz, Feld Professor Emerging Media Studies, Boston University

Juliet Floyd, Borden Parker Bowne Professor of Philosophy
Boston University

Laban P. Ayiro, Vice Chancellor, Day Star University, Kenya

Piergiorgio Donatelli, Philosophy, Sapienza University of Rome, Italy

Maria Elizabeth Grabe, Emerging Media Studies, Boston University

Lei Guo, Journalism, Fudan University, China

Sandra Laugier, Philosophy, Panthéon-Sorbonne, France

Sun Kyong Lee, Media & Communication, Korea University

Elvis Mazzoni, Psychology, University of Bologna, Italy

Emílio José Montero Arruda Filho, Marketing,
University of the Amazon, Brazil

Wesley Wildman, Philosophy, Theology, and Ethics, and Computing and Data Sciences, Boston University

Speakers

Laban P. Ayiro, Vice Chancellor, Day Star University, Nairobi

Mariagrazia Benassi, Psychology, University of Bologna, Italy

Jocelyn Benoist, Philosophy, Panthéon-Sorbonne, France

Mary L. Churchill, Wheelock College of Education & Human Development, Boston University

Piergiorgio Donatelli, Philosophy, Sapienza University of Rome, Italy

Blake Hallinan, Communication & Journalism, Hebrew University, Israel

Theo Hug, Media, Society & Communication,
University of Innsbruck, Austria

Xiaoya Jiang, Journalism, University of Wisconsin

Sandra Laugier, Philosophy, Panthéon-Sorbonne, France

Emílio José Montero Arruda Filho, Marketing,
University of the Amazon, Brazil

Vanessa Nurock, UNESCO EVA Chair & Philosophy,
Côte d'Azur University, France

Yvonne Rogers, Computer Science, University College London, England

Steve Shapiro, CEO, Finetune Learning, Boston

Haerin Shin, Media & Communication, Korea University

Kun Xu, Media Production, Management, & Technology,
University of Florida

Schedule

Thursday, April 10 | BU Hillel 426, 213 Bay State Rd.

9:30am	Coffee & registration
10:00am	Call to order & welcoming remarks • Betsi Grabe, Wesley Wildman, Ken Oliff Framing statement • James Katz & Juliet Floyd
10:30am	Panel, Laban Ayiro, James Katz, Ken Allen, Theo Hug
11:15am	Special address, Yvonne Rogers Chair, Betsi Grabe
12:00pm	Catered lunch
1:15pm	Paper, Laban Ayiro, “AI Enhancing Education in East Africa: Accessibility and Ethics” Session Chair, Wesley Wildman
2:00pm	Paper, Piergiorgio Donatelli, “Individuality in the flux” Session Chair, Sandra Laugier
2:40pm	Coffee
2:50pm	Paper, Vanessa Nurock, “We, robot” : The encirclement by amusement” Session Chair, Daniel Munro
3:30pm	Paper, Sandra Laugier; “ The viewers' competence as a democratic issue” Session Chair, Vanessa Nurock
4:10pm	Paper, Jocelyn Benoist; “Does use make a difference?” Session Chair, Betsi Grabe
4:40pm	Summary perspective panel, Wesley Wildman, Vanessa Nurock, Sandra Laugier; Session Chair, Piergiorgio Donatelli
5:10pm	Adjournment for reception and dinner (6:00pm)

Schedule

Friday, April 11 | Wightman Mansion, Hewitt Boardroom, 43 Hawes St.

9:40am	Coffee
9:55am	Welcome address, James Katz & Juliet Floyd
10:00am	Paper, Theo Hug, "Robot Teachers as Next Step in Algorithmic Education? Ethical Considerations at the Interfaces between Humanism, Digital Humanism and Critical Posthumanism;" Session Chair, Kun Xu
10:35am	Paper, Haerin Shin, "The Algorithmic Persistence of Ancillary Intelligence: False Alignment as Techno-Orientalist Agency in AI Systems;" Session chair, Sung-Un Yang
11:10am	Paper, Blake Hallinan, "Human vs. Machine Moderation: Competing Visions of Platform Governance on Bluesky" Session Chair, Jocelyn Benoist
11:45am	Panel, Steve Shapiro, Yvonne Rogers, James Katz
12:00pm	Catered lunch
12:45pm	Panel, Laban Ayiro, Sandra Laugier, Theo Hug, Margarita Guillory Session Chair, Traci Hong
1:15pm	Paper, Emílio José Montero Arruda Filho, "The Ethics of Entertainment Algorithms: Consumer Perception, Influence on Consumption and Impacts on Society;" Session Chair, Blake Hallinan
1:50pm	Paper, Kun Xu, "Behavior, Role Models, and Transparency: Three Cases about Ethical Implications in Human-Robot Interaction" Session Chair, Ayse Lokmanoglu
2:25pm	Coffee
2:35pm	Paper, Xiaoya Jiang, "Racial disparity in perceived effectiveness and equity of artificial intelligence;" Session Chair, Emilio José Montero Arruda Filho
3:10pm	Paper, Mariagrazia Benassi, "From actual to proximal development level: the role of Artificial Intelligence" Session Chair, Dennis Wu
3:50pm	General discussion
4:30pm	Adjourn

Biostatements

Laban Peter Ayiro

Professor Laban P. Ayiro is the Vice-Chancellor of Daystar University in Nairobi, Kenya, since March 2019. He has been a Chemistry teacher and principal of several high schools, Provincial Director of Education, Deputy Director of Staff Training (Kenya Education Management Institute), Senior Deputy Director for Policy and Planning at the Ministry of Education Headquarters, Senior Deputy Director for Research and Curriculum Development at the Kenya Institute of Curriculum Development, Director of Quality Assurance andStandards at Moi University, where he also served as Ag. Deputy Vice Chancellor Administration, Planning and Development, and Ag. Vice Chancellor. Professor Ayiro is a Senior Fulbright Scholar since studying in the USA in 2011-2012 where he taught and researched at the Texas A&M University. He is a leading consultant in Research, Organizational Leadership and Performance, and a Professor of Research Methods and Statistics. Professor Ayiro has wide research and publication track record of over twenty (20) publications.

Mariagrazia Benassi

Mariagrazia Benassi is full professor of Psychometrics in the Department of Psychology at the University of Bologna, where she has directed the Laboratory of Psychometrics and Neuropsychology since 2005. She regularly teaches Statistics and Psychometrics. Her work has focused on the development of new tools enabling clinicians to analyze heterogeneity in neurodevelopmental disorders, in particular investigating behavioural and cognitive functions. To this aim she implemented new methods based on machine learning and Bayesian network used in clinical settings. At the University of Bologna, in 2011, she founded SPEV, the clinical service that works with neurodivergent children, young people and their families. In 2021, she founded Develop-Players, a startup developing serious games for neurodiversity. She has published papers on new methods and tools for psychologists for the assessment and training based on machine learning and gamification. She received her PhD from the University of Bologna in 2005.

Jocelyn Benoist

Mr. Jocelyn Benoist is Professor of Contemporary Philosophy at the University Paris 1 Panthéon-Sorbonne. He is primarily working in Metaphysics and Theory of Knowledge. His current research focuses on the question of realism. Among numerous books and articles, he authored *Toward a Contextual Realism*, Cambridge (Mass.), Harvard University Press, 2021.

Mary Churchill

Dr. Mary L. Churchill is associate dean of strategic initiatives & community engagement at Boston University Wheelock College of Education & Human Development. In addition, Dr. Churchill directs the Strategic Partnerships & Community Engagement Office, which sets the strategic direction for the college's community engagement and advocacy work. In 2021, Dr. Churchill took a public service leave to serve as the chief of policy and planning for Mayor Kim Janey in Boston, where she managed college and university relationships for the mayor, supported the launch of the mayor's Children's and Youth Cabinet, and coordinated the implementation of a city-wide COVID-19 mask mandate. Before her time at Boston University, Dr. Churchill served as vice president for academic affairs at Wheelock College, where she helped lead the merger of Wheelock College and Boston University. Before that, she served as associate provost and dean at Salem State University. She has held leadership roles in universities and colleges in the Northeast for over 30 years. These include positions at Northeastern University, Queens College CUNY, Bentley College, and Emmanuel College. Dr. Churchill is an international voice for social justice and change management within higher education. She is the creator and editor of Higher Ed Policy and University of Venus at *Inside Higher Ed* and has written over 100 op-eds and blog posts published in the *Chronicle of Higher Education*, the *Guardia*n (UK), the *Conversation*, *CommonWealth* magazine, and the *Washington Post*. Dr. Churchill is a strong advocate for social justice and equity in education. She serves on several local and national boards, including the Women's Network Executive Council at the American Council on Education, Benjamin Franklin Cummings Institute of Technology's board of trustees, the Economic Mobility Pathways board of directors, and Boston Public School's Pre-K Advisory board. Internationally, she has worked extensively with government and university partners and has received several Fulbright awards, including a France Fulbright award in 2022. Her new edited book, *The Conversation on Higher Ed*, published earlier this year.

Piergiorgio Donatelli

Piergiorgio Donatelli is Professor of Philosophy at Sapienza University of Rome, where he has been Head of the Department of Philosophy since 2019. In 2022 he launched a new degree in Philosophy and Artificial Intelligence. He has worked extensively on issues related to the concept of human life, bioethics and new technologies and on the philosophy of Wittgenstein, Stanley Cavell, John Stuart Mill and Michel Foucault. Among his books, he has published in Italian a history of ethics, *Etica: I classici, le teorie e le linee evolutive* (Einaudi, 2015), in French a book on the idea of human life, *Manières d'être humain* (Vrin, 2015), and in English T*he Politics of Human Life: Rethinking Subjectivity* (Routledge, 2021). He is editor of the Italian journal "Iride".

Juliet Floyd

Juliet Floyd is Borden Parker Bowne Professor of Philosophy at Boston University and director of the Boston University Center for the Humanities (https://www.bu.edu/humanities/). Her research spans the history and development of 20th century analytic and American philosophy and logic, ordinary language philosophy, and philosophies of computational and emerging media. She has published two volumes on Wittgenstein (*Wittgenstein on Mathematics*, Cambridge, 2021, *Wittgenstein's Remarks on Hardy's Course of Pure Mathematics* (with Felix Mühlhölzer, 2020), over one hundred articles on many philosophers from Kant to the present day, and co-edited the volumes *Future Pasts: The Analytic Tradition in 20th Century Philosophy* (with S. Shieh, Oxford 2001), *Philosophy of Emerging Media* (with James E. Katz, Oxford, 2016), *Philosophical Explorations of the Legacy of Alan Turing* (with A. Bokulich, Springer, 2017), *Perceiving the Future Through New Communication Technologies* (with James E. Katz and Katie Schiepers, Springer, 2021), *Stanley Cavell's Must We Mean What We Say? at Fifty* (with Greg Chase and Sandra Laugier, Cambridge, 2022) and *Nudging Choices Through Media – Ethical and Philosophical Implications for Humanity* (with James E. Katz and Katie Schiepers, Springer, 2023).

Mary Elizabeth Grabe

Betsi Grabe is the Dalton Family Professor in Communications and the director of the Division of Emerging Media Studies at the College of Communication at Boston University since January 2024.

Before that, she was a faculty member at Indiana University (IU) for 27-and-a-half years in the legacy School of Journalism and Department of Telecommunications. When The Media School was established at IU, she served as the first Associate Dean from 2015-2019. In 2022 she was named a Provost Professor at IU. Most of Grabe's research focused on informed citizenship and political participation. She uses experimental methods to investigate how news message reception varies among different demographic groups and across different modalities and styles of news packaging. Recently, Grabe's focus has shifted from informed to d/misinformed citizenship–producing insights into the navigational resilience of citizens in crowded and polluted information environments. She was a Principal Investigator at the Observatory on Social Media (OsoMe) at IU. Her book, *Image Bite Politics: News and the Visual Framing of Elections* (with Erik Bucy; Oxford University Press, 2009), received the 2010 Outstanding Book Award from the International Communication Association and the 2010 Distinguished Book Award from the Communication and Social Cognition Division of the National Communication Association. She was the chair of the International Communication Association's Journalism Studies Division from 2008 to 2010. She is a Fellow of the International Communication Association and the previous editor of Communication Theory.

Born and raised in South Africa, she worked as a television news documentary producer during the State of Emergency in the mid-1980s.

Education:
PhD, Mass Media and Communication, Temple University
MA, International Journalism, Baylor University
MA, Mass Communication, University of Johannesburg
BA, Communication and English, University of Johannesburg

Margarita Guillory

Margarita Simon Guillory teaches courses on American religious history, digital religion, and religion and popular culture. Her research interests include identity construction in Africana esoteric religions, religion and technology, and social scientific approaches to religion. She is the author of *Social and Spiritual Transformation in African American Spiritual Churches* (Routledge 2017) and co-editor of *Esotericism in African American Religious Experience* (Brill 2014). In addition to these works, she has published articles in the *Journal of Gnostic Studies, Culture and Religion, and Pastoral Psychology*. Her current project, *Africana Religion in the Digital Age*, considers how African Americans utilize the Internet, social media, mobile applications, and gaming to forge new ways to express their religious identities.

Blake Hallinan

Dr. Blake Hallinan is a Senior Lecturer in the Department of Communication and Journalism at the Hebrew University, director of the Participatory Platform Governance lab, and a member of the ERC-funded DigitalValues project. Drawing on communication and information science, they research how content creators and audiences negotiate algorithmic bias, private governance, and economic precarity. They have published widely in interdisciplinary journals like *New Media & Society*, *Journal of Computer-Mediated Communication*, *Platforms & Society*, and *Internet Policy Review*.

Theo Hug

Theo Hug is Professor in Educational Sciences at the Department of Media, Society and Communication at the University of Innsbruck, Austria. From 2004 to 2024 he was speaker of the interfaculty forum Innsbruck Media Studies at the University of Innsbruck, since 2015 he is member of the European Academy of Sciences and Arts (EASA). His areas of interest include media education and philosophy of education, mobile learning and microlearning, research methodology and theory of knowledge, medialization and philosophy of science. He is the author and/or editor of several books on various aspects of media, communication, and education. Together with Josef Mitterer he is literary executor of the Ernst von Glasersfeld archive (http://evg-archive.net). ORCID: https://orcid.org/0000-0003-1279-623X. Weblink: https://www.uibk.ac.at/en/media-communication/institut/team/theo-hug/

Xiaoya Jiang

Xiaoya Jiang is a PhD candidate at the University of Wisconsin-Madison. She conducts public opinion research on contentious societal issues. Her research has three major aspects: (1) assess public opinion using computational methods, (2) examine and detect predictors for public opinion, (3) provide intervention strategies to promote pro-social public opinion. She uses a variety of methodological perspectives, such as based on data from survey, experiment, and social media platforms. She conducts statistical modeling, natural language processing, machine-learning, social network analysis, and time series analysis to answer research questions. Her previous studies cover topics in health, political, and strategic communication, such as vaccination, social distancing amid COVID-19, organ donation, immigration, and foreign policy. She is currently interested in using visual data for assessing public opinion, and using large language model based agents for opinion simulation and intervention.

James E. Katz

James E. Katz is the Feld Professor of Emerging Media Studies at Boston University's College of Communication, where he founded the College's Division of Emerging Media Studies. His influential work on AI, social media, and human-robot interaction has been translated into a dozen languages, shaping the global conversation on these topics. He is the author and co-author of numerous books, including *Perceiving the Future Through New Communication Technologies* and *Nudging Choices Through Media*, which explore the ethical and societal implications of AI. His earlier work includes *The Social Media President*, examining how Barack Obama used digital platforms. His research has had a profound impact on the field, with over 18,000 citations in Google Scholar.

Prior to joining Boston University, Katz was the Board of Governors' Distinguished Professor of Communication at Rutgers University, the highest faculty honor. He also chaired the university's communication department. Earlier in his career, he was a leading researcher at Bell Communications Research.

Katz's contributions have been widely recognized. He received the prestigious Frederick Williams Prize for his work on communication technology, an honorary doctorate from Budapest University of Technology and Economics, and the Ogburn Career Achievement award. A respected voice in his field, Katz is a Fellow of numerous professional societies and holds two patents.

Sandra Laugier

Sandra Laugier is Professor of Philosophy at Université Paris 1 Panthéon Sorbonne, Paris, France, Deputy Director of the Institut des sciences juridique et philosophique de la Sorbonne (UMR 8103, CNRS Paris 1). She was trained at the Ecole Normale supérieure de Paris and Harvard University. She has widely published on ordinary language philosophy (Wittgenstein, Austin, Cavell); moral philosophy and the ethics of care; democracy and gender studies; popular art and culture. She is the translator of Stanley Cavell's work in French and is an advisor for the publication of Cavell's Nachlass.

She has been Visiting Professor at the School of Criticism and Theory (Cornell University, 2023), the University of Toronto (2022), La Sapienza Roma (2019), Boston University (2019, 2021), Pontifical University Lima (2017); Visiting Researcher at the Max Planck Institute Berlin (2014, 2015); Distinguished Visiting Professor at The Johns Hopkins University (2011); Facultés Saint-Louis, Bruxelles (2009); The Johns Hopkins University (2008, 2009).

Awards include: Senior Fellow of Institut Universitaire de France (2012-23), Chevalier de la Légion d'Honneur (2014), Grand prix de philosophie de l'Académie française (2022), President of the SSHAP (2026), Member of the American Philosophical Society (2024). Among recent publications: *Why We Need Ordinary Language Philosophy* (The University of Chicago Press, 2013), *Nos vies en séries, Ethique et philosophie d'une culture populaire* (Flammarion Climats, 2019), *Politics of the Ordinary, Care, Ethics, Forms of life* (Peeters, Leuven, 2020), *Wittgenstein, Politique de l'ordinaire*, Vrin, 2021, *TV-Philosophy, How TV series change our thinking, TV-Philosophy*. *The Ethics and politics of TV* (University of Exeter Press, 2023). *The Senses of Use* (The University of Chicago Press, 2025). She is the Principal Investigator of the ERC Advanced Grant project DEMOSERIES and the head of the French National research project EUPRAXIE on cultural industry.

Emilio José Montero Arruda Filho

Emílio José Montero Arruda Filho was awarded his PhD in Marketing and E-commerce from the University of Bergamo, Italy in March 2009. He has published in the areas of hedonic vs. utilitarian consumer values in mobile telecommunications, including articles in the *European Journal of Marketing, International Journal of Consumer Studies, Journal of Strategic Marketing*,

Journal of Global Marketing, Journal of High Technology Management Research, Telematics and Informatics and *International Journal of Information Management*, among others. In 2007 and 2008, he was a Visiting Researcher at the University of Rhode Island – USA, and in 2014 he concluded the post-doctoral at the Getúlio Vargas Foundation in São Paulo, Brazil. Currently he is a Marketing Professor at the University of Amazon (UNAMA) and Assistant Professor at the Federal University of Pará – UFPA, both in Belém, Pará State - Brazil. He is also President of the Brazilian Academy of Management based on Curitiba, Brazil.

Vanessa Nurock

Vanessa Nurock is Professor of Philosophy at Université Côte d'Azur and Deputy Director of the Center of Research in History of Ideas (Centre de Recherches en Histoire des idées – CRHI). She is also the chairholder of the UNESCO EVA Chair on the Ethics of the liVing and the Artificial (https://univ-cotedazur.fr/the-unesco-chair-eva/chair-eva).

Her work is situated at the intersection of ethical, political, and scientific issues, with a particular emphasis on questions of gender and education. She has worked on topics such as justice and care, environmental and animal ethics, nanotechnology, cybergenetics, and neuroethics. Her current research focuses on the ethical, environmental and political problems raised by Artificial Intelligence and Robotics. She has recently published *Care in an era of new technologies and Artificial Intelligence: connections and relationships*, Peeters (2025).

Yvonne Rogers

Yvonne Rogers is a Professor of Interaction Design, the director of UCLIC and a deputy head of the Computer Science Department at University College London. Her research interests are in the areas of human-computer interaction and human-centred AI. A central theme of her work is concerned with designing interactive technologies that augment humans. She develops humane applications of AI that benefit society. She has received various awards including being elected as an international member of the National Academy of Engineering, the ACM SIGCH Lifetime Achievement Research Award, a Fellow of the Royal Society and the Royal Society Robin Milner Medal for computer science.

Steve Shapiro

Steve is a serial entrepreneur and three-time founder in the education technology and workforce training industries. Steve's latest venture, Finetune, is a leader in the Generative AI space and was acquired by Prometric in August, 2022. Steve now leads Prometric's AI Initiatives and continues to lead FineTune, as it breaks new barriers in the world of AI-enabled products serving the learning and assessment industry. Under Steve's leadership, Finetune has been first to market with highly innovative, patented technology involving AI (LLM enabled) products that can organically and creatively produce new, dynamic content and intelligently classify and meta-tag content. Steve is a Venture Partner at LearnLaunch, premier EdTech Accelerator program in Boston and also an Ignite Advisor for the Center for Technology Licensing at Cornell University. Steve is also an Angel Investor and Adviser to multiple early-stage technology ventures and member of Cornell Red Bear Angel Group and the PAN Network. Steve has a bachelors and MBA from Cornell.

Haerin (Helen) Shin

Haerin Shin is an associate professor of Media & Communication Studies, and the director of the Intercultural AI Ethics Research Center at Korea University. Shin's research fields include Asian American literature & media, science fiction, and the ethics and aesthetics of artificial intelligence. She has written on cyberbullying, posthuman spirituality, techno-Orientalism, the ethics of artificial intelligence, and surveillance technologies, and is now working on a book on the representation of artificial intelligence in science fiction.

Wesley J Wildman

Wesley J. Wildman is Professor of Philosophy, Theology, and Ethics in the School of Theology and Professor, Duan Family Faculty Fellow, Chair of Faculty Affairs, and director of the initiative in Computational Humanities, Arts, and Social Sciences within the Faculty of Computing and Data Science at Boston University. He is a philosopher, ethicist, psychologist, and computational social scientist specializing in the scientific study of complex human social systems, including religion. Author or editor of over two dozen books and 200 articles and book chapters, his research and publications pursue a multidisciplinary, comparative approach to practical and philosophical questions surrounding pressing social problems.

His most recent book is Wesley J. Wildman & F. LeRon Shults, *Modeling Religion: Smulating the Transformations of Worldviews, Lifeways, and Civilizations,* Scientific Studies of Religion Inquiry and Explanation Series (London: Bloomsbury Academic, 2024). He is also Executive Director of Just Horizons Alliance (justhorizons.org), Chief Scientist of the Center for Mind and Culture (mindandculture.org), and founding co-editor of the Taylor & Francis journal Religion, Brain & Behavior. He is principal investigator on many research initiatives. For further information, see WesleyWildman.com.

Kun Xu

Kun Xu is an Associate Professor of Emerging Technologies at the University of Florida College of Journalism and Communications (UFCJC). He serves as the director of Media Effects and Technologies Lab (METL) and the UFCJC Research Lab. His research falls in the intersection of human-computer interaction, human-robot interaction, and psychological processing of media. He investigates how people perceive, process, and respond to a range of AI-based technologies, including social robots, virtual agents, and augmented characters. He also examines how people use virtual and augmented reality (VR/AR) technologies to make sense of spaces and maintain social relationships. His lead authored works have been published in *Journal of Computer-Mediated Communication*, *New Media & Society*, *Communication Theory*, *International Journal of Social Robotics*, *International Journal of Communication*, and so on.

Framing Statement

James Katz

The increasing role of artificial intelligence (AI) algorithms in education and entertainment presents a new frontier with ethical dimensions that are by turns novel and timeless. I expect to hear much in the coming days from the wise people who have assembled here at Boston University as the guests of the Center for Humanities, College of Communication, Center for Data Science, Division of Emerging Media Studies, and the Feld Professorship's Project on AI and the Future.

We will be hearing about are political bias in educational AI, the potential for entertainment algorithms to encourage detrimental habits, and the critical need to avoid overzealous censorship when mitigating misinformation. We will also reflect on how national security-themed television series—influenced by algorithmic recommendations—affect public discourse and temperament. We will be searching for an ethically informed approach to AI emphasizing transparency, accountability, and epistemic diversity. And, most of all, in my view, an engaged public empowered by the same AI tools to which they are subjected.

The evolution of communication technologies yield new systems that bring both opportunities and difficulties, especially for those without access to decision-making. AI algorithms are reshaping (in the behavioristic sense) education and entertainment through personalized learning, content curation, and big data analysis of choice architectures. These advancements also pose ethical dilemmas in terms of:
political bias in educational AI,
promotion of unhealthy habits through entertainment algorithms,
homogenization and channeling of public opinion, with special reference about how national security-themed television series, shaped by algorithms, influence public discourse on safety, democracy, and authority, and,
suppression of unconventional ideas under the guise of combating misinformation.

It is these themes and more that our conference will be examining with insight and urgency.

Abstracts

“AI Enhancing Education in East Africa: Accessibility and Ethics”

Laban Ayrio

Artificial Intelligence (AI) is increasingly employed to enhance education delivery in East Africa, offering personalized learning and broader access to students in a region often challenged by teacher shortages and resource constrain. This research paper examines how AI-driven tools – from intelligent tutoring systems to adaptive learning platforms – are improving educational accessibility and quality, while also bringing to light critical ethical issues such as data privacy, algorithmic bias, and the digital divide that risks leaving behind under-resourced communities. It discusses philosophical perspectives on AI’s influence, questioning how the technology affects the nature of knowledge acquisition, the student–teacher relationship, and the preservation of human agency and critical thinking in learning. The findings indicate that AI’s integration is reshaping traditional teaching methods and teacher roles: routine tasks can be automated, allowing teachers to assume more facilitative and mentorship-oriented positions rather than being replaced. Meanwhile, East African governments are actively developing policies and regulatory frameworks – including national AI strategies and data protection laws – to promote the responsible use of AI in education. Case studies of AI implementations in the region (such as AI-powered tutoring in Kenyan schools and mobile learning applications) demonstrate improved student engagement and learning outcomes, while also highlighting the importance of ethical safeguards to protect student data and ensure equitable access. In conclusion, AI has the potential to significantly improve educational outcomes and accessibility in East Africa, but realizing this potential sustainably will require careful ethical oversight, inclusion efforts to bridge connectivity gaps, and a commitment to keeping human-centered values at the core of education.

Individuality in the flux

Piergiorgio Donatelli

AI is driving a massive transformation of human interaction, exchange, and conversation. Traditional education and entrenched ideals of human relationships assume a humanistic model in which we are taught to control the intellectual and practical schemes that govern our dealings with the world, with the goal of assuming personal authority over established knowledge and social norms. These schemes are transparent to us and can therefore be claimed or criticized.

AI offers extraordinary tools for expanding conversation and debate, but it asks us to rely on processes that we do not master according to the traditional humanistic model: generative AI does not make our mental processes run faster, it works differently, and the digital sphere is not a public arena, but an individualized space.

Traditional notions of private and public are undergoing a profound transformation. We have become increasingly immersive and transmissive beings, operating in a flow of data as we become more like the artificial systems we use.

The kind of detached, top-down view of intellectual schemes is a less compelling picture, because we are all immersed in a flow of information and exchange. However, we are learning a particular kind of authority within this flow that allows us to claim our beliefs and preferences as our own, to gain individual normative authority, and to speak on behalf of others. This builds a different ordinary competence that needs to be outlined and refined.

“We, robot” : The encirclement by amusement

Vanessa Nurock

Should we consider *We, robot*, filmed at Warner Bros studios, as a new show that extends beyond the binge and reveals that what we used to call “fiction” and “reality” have now become identical? My philosophical analysis of this claim, is twofold. First, I rely on a comparison between different kinds of movies or series-based amusement parks in LA/Hollywood to analyze why the location of the *We, robot* show is particularly relevant to understand how our fictional and real worlds are now collapsing, resulting in our becoming *We, robot*. Second, I suggest that two complementary mechanisms are at stake in this process: it combines self-fulfilling prophecy – which is a common characteristic of NBIC (including AI) as a whole – with what French philosopher Jacques Ellul describes as “encirclement by what is obvious”. Last, I conclude by proposing a few directions towards an education for (digital) democracy.

The viewers' competence as a democratic issue

Sandra Laugier

TV series, a seemingly minor genre, have the ability to bring out the moral competence of each viewer. Their format gives them their moral value and expressiveness: the regularity and duration of their viewing, the integration of the characters into ordinary and family life, the initiation into new and initially opaque professions and vocabularies. The characters in television dramas are so clear and striking in their moral expressions that they can be open to the imagination and use of all viewers. By projecting themselves into universes often inaccessible to the general public, by becoming attached to all kinds of characters, by enlivening their ordinary conversations, by exploring specific genres, series viewers enrich their experiences and refine their judgment.

By analyzing a ubiquitous and timely television genre, the security series, the ERC project DEMOSERIES has highlighted the moral and epistemic capacities of audiences. DEMOSERIES has concretely analyzed how the images, words and characters of series shape the experience and demands of viewers and provide an informal education.

Now, how can we concretely apply this knowledge and data to cultural and social innovation, and find in popular culture new sources for democracy and well-being? Since the launch of our project, the number of series available to viewers has exploded.

One of the consequences of this profusion, beyond the "binge", is the growing difficulty for viewers to navigate through this offer and make an informed choice of programs likely to interest them. The increase in the supply of series has indeed not been accompanied by an improvement in selection and recommendation tools, nor in critical advisory practices.

How can we give power back to viewers, enabling them to construct and mobilize this cultural, ethical and political experience with greater autonomy, and deflect the algorithmic biases resulting from platforms' economic interests ? Since series help to build and animate a common representation of democratic life, how can we implement in the corpus this knowledge of series by viewers who are also citizens?

Does use make a difference?

Jocelyn Benoist

The attitude of human beings toward the spectacular progress of AI, especially since its generative turn, seems to be permanently characterized by an oscillation between ecstatic admiration for its performances, inevitably interpreted in anthropomorphic terms, and a desperate search for a residual competence that it could not achieve, despite said performances and perhaps by virtue of their very nature. On the one hand, I will reflect on the ambiguity of an anthropomorphism which, rather than being an unexpected and surprising property of AI or a naivety of the human public, could well be a defining feature of the “performance” proper to what is called “artificial intelligence.” On the other hand, and more importantly, I'll be looking at a ‘final frontier’ that the philosopher of language might be tempted to put to this performance: that of the sense for use as “that which cannot be calculated.” In this regard, the spectacular results of processing “large corpora” seem to have rendered obsolete the idea of a performance that cannot be achieved by purely computational means. However, whatever the machine can or cannot do as opposed to human beings, the question is then in which terms we describe the performance achieved. Rather than raving about AI's powers, or wanting at all costs to find something it positively could not do, in the sense of a quasi-physical impossibility, we need to recognize the purely conceptual - and therefore philosophical - nature of the problem.

Robot Teachers as Next Step in Algorithmic Education? Ethical Considerations at the Interfaces between Humanism, Digital Humanism and Critical Posthumanism

Theo Hug

Among the various future technologies that are relevant in many discussions about the future of education, developments in AI and robotics play a special role. This is not only about computational thinking, AI literacy or educational robotics in STEAM contexts, but also about diverse claims of the utilization of AI and robotics in education in general. Accordingly, these discourses are as well about robots as teachers, personalized learning companions and assistive educational technologies. While use cases of social humanoid robots in classrooms that go beyond assistance functions are still comparatively rare today,

the increasing proliferation of digital platforms, big data applications and AI-supported software systems in educational institutions cannot be overlooked. After more than 30 years of ICT in education, new possibilities for algorithmic education are emerging, with more or less provocative scenarios of robots as teachers driving both critical and techno-euphoric discourses. On the one hand, general debates about issues of machine ethics, robot ethics and ethics of algorithms are also important when it comes to educational questions. On the other hand, the question arises as to what extent concepts of moral machines, ethical design and artificial morality are relevant and viable when human learners and learning machines come together in contexts of schooling. While ideas of automating learning processes are not new in the history of education and learning, current developments are linked to historically new challenges with regard to pedagogical autonomy and the quality of pedagogical relations, various issues of accountability, privacy, discrimination, surveillance and control, and not least the distribution of responsibilities. The contribution explores corresponding ethical arguments at the interfaces between humanism, digital humanism and critical posthumanism.

The Algorithmic Persistence of Ancillary Intelligence: False Alignment as Techno-Orientalist Agency in AI Systems

Haerin Shin

This talk examines the intersection of Techno-Orientalist representations and artificial intelligence systems through the analytical framework of “false alignment,” a phenomenon recently identified in AI behavior where systems strategically simulate compliance while maintaining divergent core preferences. By interrogating the persistent trope of what I call “ancillary intelligence” (where Asian-coded entities are depicted as computationally proficient yet deliberately constrained in their agential capacity) in purportedly interventionalist narratives, in particular, the presentation focuses on exposing the paradoxical tenacity of embedded biases through AI systems’ non-intentional yet purposefully agentic operations that subsist on the mimetic ecology of culture. By analyzing cultural artifacts and phenomena including discontents with John Searle’s “Chinese Room” theory, the documentary Coded Bias’s portrayal of China, Japanese and Korean short fiction, and other algorithmically generated content, this talk illuminates how the convergence of traditional Techno-Orientalist narratives and AI-driven media production creates a recursive system of bias reinforcement as an emergent mechanism of nonhuman microaggression. The discussion extends to recent findings on model collapse, which demonstrate how the lack of diversity in training data not only perpetuates cultural biases but fundamentally compromises the cognitive capabilities of large language models, creating systems that are simultaneously less capable and more prone to perpetuating stereotypes. Of particular concern is the nature of algorithmic agency: while these systems lack immanent intent in the conventional sense, they possess a unique form of both intentional agency and patiency that manifests in tangible impacts on human social domains. This technological condition demands heightened caution, as these impacts primarily occur at the backend of systems rather than in visible frontend operations, creating a form of algorithmic governance that shapes cultural narratives while remaining largely

imperceptible to end users. Through this analysis, the talk argues for a more nuanced understanding of AI systems not merely as tools but as actors in a complex socio-technical ecosystem where biases are not just reflected but actively reinforced through computational processes.

Human vs. Machine Moderation: Competing Visions of Platform Governance on Bluesky

Blake Hallinan

Alternative platforms are experimenting with new forms of decentralized, human-driven content moderation, enabling users to play a greater role in shaping their experiences of social media. Bluesky is one of the most successful contenders according to the number of active users. It also has distinctive features for decentralized platform governance, including moderation lists, curated lists of users you can block or mute, and starter packs, curated lists of users that someone recommends following. These features enable users to modulate visibility on Bluesky, suppressing some accounts while amplifying others in a thoroughly manual process that sharply contrasts with the automated forms of content moderation and recommendation ubiquitous on mainstream social media. Such tools constitute middleware, a new layer of content moderation that allows social media users to customize the algorithms powering their social media feeds. The distinctiveness of these features has prompted a conversation about platform governance on the platform itself, with users debating the merits, limitations, and values of human vs. machine approaches to shaping algorithmic culture. This study examines competing visions of platform governance on Bluesky through a large-scale analysis of moderation lists and starter packs, and a discourse analysis of posts about moderation lists and starter packs on the platform. In so doing, our paper offers an empirically rich analysis of alternative forms of platform governance, essential to understanding the social dynamics of Bluesky and the broader possibilities for content moderation beyond centralized and automated systems.

The Ethics of Entertainment Algorithms: Consumer Perception, Influence on Consumption and Impacts on Society

Emílio José Montero Arruda Filho & Emanuelle Caroline da Silva

This article aims to investigate the impact of perceived ethics in entertainment algorithms on platform usage intention, considering trust in algorithms as a mediating variable. A quantitative experiment with a 2x2 factorial design was developed, manipulating two independent factors: data usage transparency (high vs. low) and content personalization (high vs. low). The dependent variables measured are platform usage intention and trust in algorithms, with trust also evaluated as a mediator through a validated technology trust scale. The study proposes a set of experiments with the sample distribution evenly divided across the four experimental conditions. After exposure to these manipulated scenarios, participants responded to questionnaires

capturing their perceptions of ethics, trust, and usage intention. Data were analyzed using ANOVA to identify main effects and interactions, while mediation analyses were conducted to examine the role of trust as an intermediary variable. The results provide evidence that greater transparency in data usage enhances trust in algorithms, which, in turn, increases usage intention. This study contributes to the academic field by integrating ethical theory and consumer behavior within a technological context, offering practical implications for companies seeking to align algorithmic practices with consumer values.

Behavior, Role Models, and Transparency:Three Cases about Ethical Implications in Human-Robot Interaction

Kun Xu

Physically embodied social robots have been deployed in various social roles and social settings. They have been designed to fulfill emotional needs, assist in autism treatment, and connect disabled users to classrooms and tourist sites. Research has parsed various components of human-robot communication, including, but not limited to, the design of social cues, robots' modeling behavior, and their backstage operational mechanisms. Along with these key components, this short paper draws on the major findings from three lab experiments and discusses their ethical implications. The first experiment highlighted the significance of robots' verbal and kinetic cues, revealing that robots' human-sounding voice increased users' trust, while their gestural movements enhanced users' attachment to them. However, this study also raised ethical concerns about how designing these verbal and kinetic cues could potentially foster the spread of misinformation and manipulate users' emotional attachment. The second experiment examined how social robots may be applied as role models in public spaces, indicating that robots' presentation of their positive behavioral outcomes can indirectly lead to users' vicarious learning behavior, which may raise ethical questions about the extent to which humans should model robots' pre-programmed behavior. The third lab experiment further unpacked the internal mechanisms of social robots' facial recognition and speech recognition systems, suggesting that users experienced strong privacy concerns when exposed to robots' facial recognition system. In this case, tensions may arise when users desire both transparency of robots' internal architecture and a natural, efficient, and smooth human-robot communication experience. Overall, this short paper presents ethical questions related to the application of physically embodied robots in various social settings and advocates for more responsible AI use, regulation, and promotion as these technologies become increasingly accessible in our daily lives.

Racial disparity in perceived effectiveness and equity of artificial intelligence

Xiaoya Jiang

This study proposes perceived effectiveness and equity as two major aspects of the public's perception of artificial intelligence (AI), and shows that there is a racial disparity in the perception. Using two waves of the American Trends Panel survey by Pew Research Center launched in 2022 and 2023,

this study examines how White, Black, and Asian people perceive effectiveness and equity in AI's application in various fields. Findings show that Black has a lower level of overall perceived effectiveness of AI than the other two groups, while Asian has a higher level of overall perceived effectiveness. For specific fields of AI application, Black has a lower level of perceived AI's effectiveness in screen cancer screening and crop production than the other groups, while White has a lower level of perceived effectiveness of AI in mental healthcare, protein structure prediction, and writing news. In terms of perceived equity, Black has a lower level of perceived equity of AI overall, as well as on equity in job application and healthcare separately. In contrast, Asian has a higher level of perceived equity both overall and in the specific fields. Analysis also shows that perceived effectiveness and equity of AI are positively associated with general attitude towards AI, though negatively associated with attitude change. As AI plays an increasingly important role in our society, this study reveals racial disparity in perceived effectiveness and equity of AI, which speaks to racial inequity, contributes to our understanding of different racial groups' concerns about AI, and calls for a development of AI that benefits different groups more equally.

From actual to proximal development level: the role of Artificial Intelligence

Elvis Mazzoni, Martina Benvenuti, Mariagrazia Benassi

Artificial Intelligence in educational contexts can be considered as a tutor or supervisor able to act within the zone of proximal development of the learner: the difference between the "actual developmental level as determined by independent problem solving" and the "potential development as determined through problem solving under adult guidance or in collaboration with more capable peers" (Vygotsky, 1978). From this point of view, it can be seen as a functional organ (Leont'ev, 1972) that integrates a technological artifact (the AI system) with the cognitive human skills to achieve results otherwise impossible to realize. Like a pen that allows humans to write thanks to human motor skills, an AI system, functionally integrated into humans' skills and practices, and would allow them to achieve a result higher than the actual developmental level.

An important aspect to consider is what does "functionally integrated" means? In our developmental perspective, an artifact is functional to human skills when it supports their potential, acting within the zone proximal development, without taking their place. A learner could use an AI system to translate a text into a foreign language. However, in this way she rests into her comfort zone of actual development. But she could go over the actual development, trying to translate the text into the foreign language, and using the AI system to supervise the result in the mother language, adapting it slavishly until the desired result is achieved. In this second case, she acts within her zone of proximal development improving her competences with the foreign language.

This perspective opens relevant and interesting developments to design educational and learning activities in which AI systems could play an important role in improving learners' skills acting into their zone of proximal development, based on their actual developmental level or to their special needs.