Chair of Political Economy
Prof. Nora Szech
Lixuan Zhao

Shaping AI and Digitization for Society

I. TOPIC DESCRIPTION

> "A computer would deserve to be called intelligent if it could deceive a human into believing that it was human" – Alan Turing, 1950

The pervasion of AI-based applications continues to grow at a staggering pace and digitization is also hardly a new phenomenon which continues to persist. Already today, they affect many aspects of our lives. Thus, we are confronted with the timely question of how to design AI and digital economy in ways that are optimal depending on context, users' and society's preferences.

AI and digitization promise tremendous opportunities for progress. How to harness these opportunities for the best of society, however, remains an open question. What are potential trade-offs? How can designs look like that work well according to specific measures of welfare? This is the topic of this seminar. Participants develop their own research idea on how to design AI to society's best interests, for example, with one specific and politically relevant context in mind. The goal is to come up with an experimental design (lab or field experiment) that enlightens us about some specific research question within the topic. To find such a context, an entry point may be one of the following.

Digitalization has dramatically reduced search, information and tracking costs, thus making it easier than ever for consumers to compare prices and product features. On the other hand, this development may contribute to information overload and 'attention wars'. What does that imply for market design? Or, consider the effect of negligible tracking costs such that firms can design highly targeted advertising or implement first-order price discrimination easily. (E.g., consumers using an Apple device may be offered higher prices than others.) What should we think about this from an economic design perspective? Should we regulate these markets, and if so, how?

The reliance on digital technology has grown tremendously during the pandemic and digital security in time of crisis has become an increasingly important issue. Research shows that there was an increase in the reports of cyber-attacks during the pandemic, and frauds related to online shopping and the hacking of social media and emails have witnessed the largest increases in the number of incidents (Buil-Gil et al, 2020). Alongside cyber-attacks, fake news and false information has also been circulated on the internet and social media since the outbreak of COVID-19. How could we use the insights from behavioral science to address this problem? And how should companies better design the digital platforms to limit the spread of misinformation? There are many more interesting questions to be explored here.

AI may help people to make better decisions, e.g. by giving nudges that promote healthier lifestyle choices and/or support long-term goals. But how should they become designed in optimal ways? Are there trade-offs? How should we measure welfare? Answers may be quite context-specific, of course, and it may make sense to focus on one specific one for a research question, e.g., how to

nudge people to save money, eat healthy, drink less… And, do people enjoy such nudging in the short run? In the long run? Are there differences depending on personal characteristics?

Of course, AI might also give consumers back control over their data, e.g. via apps scanning the usually too long to read Terms & Conditions agreements and provide. How could this be designed in a good way?

Furthermore, algorithms can create filter bubbles and political echo chambers. Do we have something to say to this from an information design point of view?

Finally, people fear that AI could contribute to the emergence of totalitarian regimes.

"What Orwell prophesied in 1984, where technology was being used to monitor, control, dictate, or what Huxley imagined we may do just by distracting ourselves without any meaning or purpose. Neither of these futures is something that we want. [...] There are unintended consequences of technology. It is up to us to ensure that some of the more dystopian scenarios don't come true.",

says Microsoft's CEO Satya Nadella. Preventing an Orwellian dystopia from becoming reality will crucially depend on the strength and role of political institutions. E.g., consider the case of San Francisco which has recently banned facial recognition technology in public places. How could AI foster security and nevertheless respect aspects of privacy? Suggestions need to be developed and empirically evaluated.

Topic suggestions include but not limited to: Privacy, Health Markets, Markets for Attention, Addiction, Fake News, Voting in the Digital Era, …

II.   SCOPE OF THE SEMINAR

Please note that no topics will be handed out. Instead, participants will develop their own research question in the broad field of Shaping AI and Digitization for Society. Using their own creative abilities, students design an economic experimental or field study that answers their research question. Students will work in groups of up to four, depending on the total number of participants. If the process of group formation is unsuccessful, participants will be assigned.

The seminar starts with an introductory meeting via Zoom on 20.04.2022. Ideas for experiments or field studies will be presented in a blocked event in end May/early June. Each presentation should last max. 20 minutes. Full attendance in all meetings is required for successful participation in the seminar. Group seminar papers of 8 – 10 pages, as well as two individual abstracts (one with 75 to 100 words and another with 150 to 200 words) are to be handed in at the end of the semester.Application will be possible via the centralized platform until 03.02.2022.

Please note that we will ask you for a brief description of the topic and research idea you are interested in, in order to ensure a well-balanced variety in the seminar. Prior attendance of the courses "Economics and Behavior" and/or "Auction and Mechanism Design" is recommended but not required. For further questions, please contact Lixuan Zhao (lixuan.zhao@kit.edu).

III.   RELATED LITERATURE

A. GENERAL INTRODUCTION

– Cassar, Friedman (2004): Economics Lab. An Intensive Course in Experimental Economics. *Routledge*
– Croson (2002): Why and how to Experiment. *University of Illinois Review*

B. SOURCES OF INSPIRATION

- *Consumer empowerment and digital markets*

– Chernev, A., Böckenholt, U., & Goodman, J. (2015). Choice overload: A conceptual review and meta-analysis. Journal of Consumer Psychology, 25(2), 333-358.
– Costa, E., & Halpern, D. (2019). The behavioural science of online harm and manipulation, and what to do about it. Retrieved from: https://www.bi.team/publications/the-behavioural-science-of-online-harmand-manipulation-and-what-to-do-about-it/
– Dertwinkel-Kalt, M., Köster, M., Sutter, M. (2019): To Buy or not to Buy? Shrouding and Partitioning of Prices in an Online Shopping Field Experiment, CESifo, working paper no.7475/2019.
– Gabaix, X. & Laibson, D. (2006). Shrouded attributes, consumer myopia, and information suppression in competitive markets. The Quarterly Journal of Economics, 121(2), 505-540.
– Goldfarb, A., & Tucker, C. (2019). Digital economics. Journal of Economic Literature, 57(1), 3-43.
– Kerschbamer, R., Neururer, D., & Sutter, M. (2019): Credence goods markets and the informational value of new media: A natural field experiment. MPI Collective Goods Discussion Paper, (2019/3).
– Mariotti, T., Schweizer, N., Szech, N., & von Wangenheim, J. (2018). Information nudges and self-control. [working paper, https://polit.econ.kit.edu/downloads/papers/MSSW_Information_Nudges.pdf]
– OECD. (2018). Personalised Pricing in the Digital Era. DAF/COMP(2018)13. Retrieved from: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DAF/COMP(2018)13&docLanguage=En
– Penczynski, S. P., Zhang, S. (2017): Disclosure of Verifiable Information under Competition: An Experimental Study. Working Paper, available at penczynski.de
– Sadoff, S., Samek, A. S., & Sprenger, C. (2015): Dynamic inconsistency in food choice: Experimental evidence from a food desert. Becker Friedman Institute for Research in Economics Working Paper, (2572821)
– Schweizer, N., & Szech, N. (2018). Optimal revelation of life-changing information. Management Science, 64(11), 5250-5262.

- Soman, D. (2015). The Last Mile: Creating Social and Economic Value from Behavioural Insights. Toronto, Canada: University of Toronto Press.

- *The influence of algorithms on political debates & biases*
- Braun, S., Dwenger, N., Kübler, D., Westkamp, A. (2014): Implementing quotas in university admissions: An experimental analysis. Games and Economic Behavior, 85, 232–251
- Larson, J., Mattu, S., Kirchner, L., & Angwin, J. (2016). How we analysed the COMPAS recidivism algorithm. ProPublica. 23 May. Retrieved from: https://www.propublica.org/article/how-we-analyzed-the-compasrecidivismalgorithm; Dressel, J. & Farid, H. (2018). The accuracy, fairness, and limits of predicting recidivism. Science Advances, 4(1), eaao5580.
- Schweizer, N., & Szech, N. (2017). Revenues and welfare in auctions with information release. Journal of Economic Theory, 170, 86-111.
- Staley, O. (2017). Harvard economist Iris Bohnet on how to design a hiring process that's fair for everyone. Quartz at Work. 17 October. Retrieved from: https://qz.com/work/1080530/harvard-economist-iris-bohnet-says-toeliminate-bias-its-easier-to-change-systems-than-change-people/ 4
- Sunstein, C. R. (2001). Echo chambers: Bush v. Gore, impeachment, and beyond. Princeton, NJ: Princeton University Press.
- Szech, N. (2011). Optimal advertising of auctions. Journal of Economic Theory, 146(6), 2596-2607.
- West, S.M., Whittaker, M. & Crawford, K. (2019). Discriminating Systems: Gender, Race and Power in AI. AI Now Institute. Retrieved from https://ainowinstitute.org/discriminatingsystems.html

- *Privacy/security*
- Beresford, A. R., Kübler, D., Preibusch, S. (2012): Unwillingness to pay for privacy: A field experiment. Economics letters, 117(1), 25-27.
- Conger, K., Fausset, R., & Kovaleski, S.F. (2019). San Francisco Bans Facial Recognition Software. New York Times, May 14, 2019. Retrieved from: https://www.nytimes.com/2019/05/14/us/facial-recognition-ban-sanfrancisco.html
- Orwell, G. (1949). Nineteen eighty-four: A novel. London: Secker & Warburg
- Schudy, S., & Utikal, V. (2017): 'You must not know about me'—On the willingness to share personal data. Journal of Economic Behavior & Organization, 141, 1-13.