Exploring the Nuances of Designing (with/for) Artificial Intelligence

Stoimenova, N.; Price, R.A.

DOI
10.1162/desi_a_00613

Publication date
2020

Document Version
Accepted author manuscript

Published in
Design Issues: history/theory/criticism

Citation (APA)
Stoimenova, N., & Price, R. A. (2020). Exploring the Nuances of Designing (with/for) Artificial Intelligence. *Design Issues: history/theory/criticism*, *36*(4), 45-55. https://doi.org/10.1162/desi_a_00613

Exploring the nuances of designing (for/with) AI

Accepted to Design Issues 36(4)

Ir. Niya Stoimenova
Delft University of Technology
Department of Design, Organisation and Strategy
n.stoimenova@tudelft.nl

Bio

ir. Niya Stoimenova is a Doctoral Candidate at Delft University of Technology, the Netherlands. Her research focus is two-fold: (1) identifying and addressing the unintended consequences of AI-powered solutions through design and (2) the creation and implementation of organizational structures that support such behaviour. As such, she is particularly interested in the possible future roles design(ers) could play and their potential impact on society.

Dr. Rebecca Price
Delft University of Technology
Department of Design, Organisation and Strategy
r.a.price@tudelft.nl

Bio

Dr Rebecca Price is an Assistant Professor of Transition Design at Delft University of Technology, Netherlands. Dr. Price works with public and private organizations to support the application of design upon complex innovation challenges. While the predominant domain of her work to date has been mobility (aviation, automotive, urban transport), her methodological research in particular holds increasing value to domains related to public health and energy transitions that stem from a socio-technical perspective of the possibilities of design.

Corresponding Author: n.stoimenova@tudelft.nl

Introduction

For all of the technology advancements since the creation of the internet, the infant stages of a true digital economy are only yet being realised. A fundamental shift in the way society operates is approaching - driven by advances in artificial intelligence (AI) (often referred to interchangeably as one of its branches – machine learning (ML)). There are complexities to the incorporation of AI into products, services and systems requiring attention of the design discipline. However, while the technical progress of AI is compelling, there is an observed comparative lack of discourse across the design discipline regarding the topic.

The aim of this article is to shed light on the deeper consequences of AI development for the design discipline. Our rationale is to avoid repeating hype-associated dialogue of utopian/dystopian technology futures. We do undertake scene-setting and review major perspectives within the AI landscape to frame this article. Our emphasis then moves to presenting and discussing technology-related developments in relation to the design movement. We identify that the problems AI-powered solutions address are becoming increasingly broad in scope and inherently wicked. Practical examples are drawn upon to illustrate this point. Yet AI-solutions remain riddled with strong biases that when poorly conceived can cause more damage than good.

A methodological gap emerges from our analysis. We identify that there are fragments of methodological readiness for AI within the design discipline, but these are yet to be connected to meet the nuances of designing with this new subject matter. Nuances such as the challenge of designing to ensure AI-powered artefacts remain safe even as utility evolves over time – both through instructional and machine learning (ML) prompted by the user's interactions and input from a broader system. This article brings together the methodological developments of thing-centred, participatory and human-computer interaction design in order to consider a new approach for designing with AI.

Setting the Scene

In 2016 an experimental vehicle developed by researchers at NVIDIA performed considerably better than its competitors (Alphabet, Uber, Tesla). Unlike other autonomous vehicles, 'it' learned how to behave on the road by watching a human instead of being provided instructions on how to drive. However, due to the standard way such algorithms are devised, the vehicle's reasoning processes were largely opaque – a mystery even its developers struggled to untangle.[1] To address the problem, Mariusz Bojarski and colleagues at NVIDIA Corporation, New York University and Google Research developed a simple method for highlighting the parts of the image the algorithm pays attention to.[2] However, the rationale behind why these parts of the image were highlighted remained largely unknown.

The Artificial Intelligence Landscape

The development of AI is expected to occur in three stages – narrow (weak) AI, general (strong) AI and Artificial Superintelligence or Intelligence Amplification.[3] Narrow AI, currently achieved and believed by many scientists to be the only possible incarnation of intelligent machines, is bound to one specific field and is incapable of performing tasks outside a preprogramed scope. For example; the way Netflix and Spotify generate recommendations; the use of chatbots to address customer inquiries, and; the way Facebook decides what to curate in the user's newsfeed (although contentiously). While some widely publicised AI implementations tackle more general tasks such as driving a car (Tesla's autopilot) or generating music (IBM's Watson), these examples are still considered a coordination of several narrow AIs.

---

1. Knight, Will. "The Dark Secret at the Heart of AI." *MIT Technology Review.* April 11, 2017.
2. Mariusz, Bojarski., Philip Yeres., Anna Choromanaska., Krzysztof Choromanski., Bernhard Firner., Lawrence Jackel., and Urse Muller. "Explaining How a Deep Neural Network Trained with End-to-End Learning Steers a Car." CoRR, (/ 2017).
3. Noessel, Christopher. *Designing Agentive technology: AI that works for people*. Rosenfeld Media, 2017.

The second speculative stage is general AI, which Shulman and Bostrom[4], researchers at the Future of Humanity Institute at Oxford University define as, ‘‘systems which match or exceed the [intelligence] of humans in virtually all domains of interest.’’ A growing number of renowned scientists, philosophers and forecasters predict general AI creation by mid-twenty-first century. It is their belief that accelerating progress in hardware, artificial intelligence, robotics, genetic engineering and nanotechnology make this timeframe achievable.[5] If achievable, the advent of general AI could eventually trigger an event called technological singularity.[6,7] At this time, many scholars and practitioners have argued that the growth of machine intelligence is likely to radically affect civilization. [8, 9,10]

Due to its controversial nature, this notion has resulted in two possible dichotomous scenarios for the future. The first predicts the emergence of artificial *superintelligent* agents possessing intelligence reaching far beyond the collective capabilities of all renowned human experts in knowledge fields, including scientific creativity, general wisdom and social skills as suggested by Bostrom and Yudkowsky.[11] This scenario could also be seen as somewhat akin to Georg Hegel’s description of the ascent of human culture to an ideal point of absolute knowing.[12] Many prominent entrepreneurs, scientists and philosophers such as Bill Gates, Stephen Hawking, Elon Musk and Sam Harris believe this scenario could lead to human extinction.

The second scenario describes the emergence of a posthuman race, evolved with amplification of human cognitive capabilities.[13] This new race would theoretically overcome existing physical and mental limitations; in the most extreme sense conquer disease, aging and even death. [14] For many, this type of forecasting is uncomfortable and even cause for cognitive dissonance toward the topic. Yet multiple companies are already working to explore this future scenario. Neuralink is working on developing a *neural lace* that is said to improve bandwidth of communication between human brains and machines; Facebook announcing interest in working on enabling people to type with their thoughts[15] and; the start-up CTRL-Labs demonstrating a prototype that allows users to interact with a machine by sending signals through their spinal column.[16] These technology futures have been confined to the realms of science fiction to date – but now confront us contentiously.

It is plausible that we may never enter the latter two stages of AI development. What is already occurring however, is the integration of narrow AIs into everyday life through new products, services and systems.[17] From the music and news such algorithms recommend to the way devices are unlocked with the user’s facial glance. Many services and devices regularly mine user behaviour and contextual data to tailor highly personalised services and experiences. An object like a fridge now understands the user's diet by sensing ‘its’ own shelf contents and may even purchase weekly groceries. Collectively, a network of home

4. Shulman, Carl., and Nick Bostrom. “How hard is artificial intelligence? Evolutionary arguments and selection effects.” *Journal of Consciousness Studies 19,* no. 7-8 (2012): 103-130.

6. Eden, Amnon H., James, Moor H., Jonny Soraker., and Eric Steinhart. *Singularity Hypotheses: A Scientific and philosophical assessment.* (Berlin: Springer, 2012).

7 Vinge, Vernor. “The coming technological singularity: How to survive in the post-human era” *Whole Earth Review,* (1993).

8. Paul, Gregory, S. and Earl Cox. *Beyond Humanity: Cyberevolution and Future Minds.* (Rockland, MA: Charles River Media, 1996).

9. Broderick, Damien. *The spike: how our lives are being transformed by rapidly advancing technologies.* (New York, NY: Tom Doherty Associates, 2002).

10. Kurzweil, Ray. *The Singularity is Near.* (London: Viking, 2010).

11. Bostrom, Nick., and Eliezer Yudkowsky. “The ethics of artificial intelligence” eds in *The Cambridge Handbook of Artificial Intelligence.* (Cambridge University Press, 2014): 316-334.

12. Hegel, Georg W.F. *The Phenomenology of Spirit.* (Oxford: Oxford University Press,1977).

13. Vinge, Vernor. “The coming technological singularity.” *Whole Earth Review,* Winter Issue (1993).

14. Harari, Noah Y. *Homo Deus: A Brief History of Tomorrow*. (Harvill Secker, 2015).

15. Statt, Nick. “Facebook is working on a way to let you type with your brain” *The Verge.* April 19, 2017.

16. Levy, Steven. 2017. “Brain-machine interface isn't sci-fi anymore” *Wired.* September 13, 2017.

17. Yang, Qian. "The Role of Design in Creating Machine-Learning-Enhanced User Experience." In *2017 AAAI Spring Symposium Series*. 2017.

appliances supports the user/s to live a fulfilling lifestyle. Identity becomes deeply enmeshed in a non-biological matrix of machines, tools, codes and semi-intelligent daily objects.[18] Such technologies are becoming less like tools and more like part of an extended mental apparatus of the person.[19] The maturity and prevalence of this technology has catalysed the notion that 'ML is the new UX'. That is, ML will be the most important way to improve user experience.[20] Such depth of knowledge initiates the transition from highly personalised experiences to personalised realities highlighting the interplay between expert and everyday ideas of appropriate practice. This is not just a practical issue, concerning 'what works' and 'what does not work' under specific circumstances. It holds profoundly ethical implications; who determines what artificial 'doings' that deeply affect how society operates can do.[21] To unpack this conundrum, we first need to understand the dual nature of AI (ML) problems, even in a narrow state.

Duality of problems in narrow AI

In April 2018 Ryen White and co-authors from Microsoft Research and Duke University published a paper reporting their initial attempts to create a "simple scalable test that can be used for screening of Parkinson's disease in the community or at home."[22] The researchers used longitudinal log data from Microsoft's search engine, Bing, to look into the presence and frequency of symptom-related query terms; motor symptoms such as the speed, direction and tremors of cursor movements, and presence of risk factors. Despite still being in a testing phase, their model showed promise in detecting a disease that has a current clinical early diagnosis accuracy of approximately 80%.[23] Similarly, Stanford University researchers devised an algorithm that performed better than radiologists in detecting pneumonia from frontal-view chest X-ray images.[24] Scientists from Google, Harvard University and Connecticut University created an algorithm that can forecast the aftershock locations of earthquakes and identify, "physical quantities that may control earthquake triggering during the most active part of the seismic cycle."[25] These examples are complemented by advances in self-driving cars and their accompanying new concepts of mobility; anti-aging efforts (e.g. Alphabet's Calico), and optimising agriculture (e.g. FarmLogs). However, while all these models are promising, the consensus within the scientific community is that these initial use cases still require further longitudinal investigation.

ML applications like these are used to automatically detect patterns in data and then use these to predict future data.[26] As such, they inherently solve what Rittel and Webber term as tame problems.[27] Tame problems involve an enumerable set of solutions, clear rules and binary decision mechanisms (true or false). However, what is devised to solve the tame problem of predicting data patterns, transcends its initial boundaries and begins to impact the larger social system in which it is situated. At this point, the tame solution enters and interacts with thewicked environment. According to Rittel and Webber, wicked problems are a "class of social system problems which are ill-formulated" and defined by confusing information, multiple clients and decision makers with conflicting values, and significant ramifications.

---

18. Clark, Andy. "Natural-born cyborgs?." eds In *Cognitive technology: Instruments of mind*. (Springer, Berlin, Heidelberg; 2001): 17-24.
19. ibid
20.Yang, Qian., John Zimmerman., and Aaron Steinfeld. "Review of Medical Decision Support Tools: Emerging opportunity for Interaction Design." In *Proceedings of IASDR 2015 Interplay,* 2015.
21. Kuijer, Lenneke., and Elisa Giaccardi. "Co-performance: Conceptualizing the role of artificial agency in the design of everyday life." In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, p. 125. ACM, 2018.
22. White, Ryen W., P. Murali Doraiswamy., and Eric Horvitz. "Detecting neurodegenerative disorders from web search signals." *Digital Medicine* 1, no. 1 (2018): 8.
23. ibid
24. Rajpurkar, Pranav., Jeremy Irvin., Kaylie Zhu., Brandon Yang., Hershel Mehta., Tony Duan., Daisy Ding et al. "Chexnet: Radiologist-Level Pneumonia Detection on Chest C-Rays with Deep Learning." *Working paper; arXiv:1711.05225v3*, Cornell University Library, (2017.
25. DeVries, Phoebe M.R., Fernanda Viégas., Martin Wattenberg., and Brendan J. Meade. "Deep learning of aftershock patterns following large earthquakes." *Nature 560*, no. 7720 (2018): 632.
26. Murphy, Kevin. *Machine Learning: A Probabilistic Perspective*. (Cambridge, Mass, MIT Press: 2012)
27. Rittel, Horst W.J., and Melvin M. Webber. "Wicked problems." *Man-made Futures 26*, no. 1 (1974): 272-280.

To exemplify this transition of AI from tame to wicked problem, we return to the work of White and colleagues. The focus of their study is accurately predicting the presence of factors that signal potential Parkinson's disease. However, their ambition to create a Parkinson's test for community and home use inevitably posits a social problem. The project encounters sensitivities concerning quality of life and mortality. The project responds to the problem formulation – *identify and diagnose*. Yet the information delivery of this type of solution requires careful consideration of the perspectives within allied health system. How does the solution deliver diagnosis? What role do doctors now play when diagnosis is outsourced? How does the solution connect the diagnosed within the allied health service? How might the family require support post-diagnosis? Consequently, there is no enumerable set of potential solutions nor a well-described set of permissible operations. Delivering false diagnosis or delivering diagnosis insensitively could cause significant distress to the community.

In this arrangement, a solution is under pressure to surpass true *or* false criteria (identify and diagnose), and move to an ethical evaluation; right *and* wrong. Yet, as Rittel and Webber note, since "many parties are equally equipped and interested to judge the solution, [sic] none has the power to set formal decisions rules to determine correctness." [28] Finally, there is no immediate and ultimate test of a solution to the problem that ensures positive impact. Every solution to this problem is a "one-shot operation" and the epitome of a wicked problem described by Rittel and Webber.

Two prominent examples highlight how a tame-problem-solving algorithm may cause distress when interacting with the broader social system. First, the algorithms Google and Facebook use to rank pages and show content. In 2016 these algorithms were used as a tool for mass misinformation and manipulation, resulting in the infamous term "fake news". However, despite both companies' efforts to prevent such situations after the US election in 2016, the spread of fake news after the mass shooting in Las Vegas in October 2017 proved very difficult to prevent.[29] Second, that of ProPublica. In 2016 ProPublica ran an experiment with an algorithm widely used within the US judicial system. The algorithm exhibited deeply-seeded racial biases and proved highly inaccurate, resulting in many falsely identified defendants, described in the work of Julia Angwin and colleagues.[30] In both cases, the ML-powered solutions were driven by data availability and learner performance rather than deliberate vision.[31] Thus, they failed to account for users in various scenarios despite being widely adopted in many newly introduced consumer goods and services. [32, 33]

James Guszcza and colleagues from Deloitte in partnership with MIT lab note; "It is by now abundantly clear that, left unchecked, AI algorithms embedded in digital and social technologies can encode societal biases, accelerate the spread of rumors and disinformation, amplify echo chambers of public opinion, hijack our attention, and even impair our mental wellbeing."[34] While Guszcza and colleagues emphasize reactive auditing of algorithms once implemented to correct wrongful activity, we see great value in the methodological developments within design that stem from a tradition of dealing with wicked problem.

Views on design with AI/ML

---

28. ibid

29. Eordogh, Fruzsina "Goolge needs to Blacklist 4 chan During National Crises" *Forbes.* October 3, 2017.
30. Angwin, Julia., Surya Mattu., and Lauren Kirchner. "Machine Bias" *ProPublica,* May 23, 2016.
31. ibid.
32. Amershi, Saleema, Maya Cakmak, William Bradley Knox, and Todd Kulesza. "Power to the people: The role of humans in interactive machine learning." *AI Magazine 35,* no. 4 (2014): 105-120.
33. Yang, Qian., John Zimmerman., Aaron Steinfeld., Lisa Carey., and James F. Antaki. "Investigating the Heart Pump Implant Decision Process: Opportunities for Decision Support Tools to Help." *ACM transactions on computer-human interaction : a publication of the Association for Computing Machinery 2016,* no.1 (2016): 4477-88.
34. Guszcza, James., Iyad Rahwan., Will Bible., Manuel Cebrian., and Vic Katyal. "Why We Need to Audit Algorithms." *Harvard Business Review,* (2018).

Discussion within the design discipline regarding the implications of AI on design is limited, with scholars from human-computer interaction (HCI) thus far taking leadership. Discourse here concerns the best paradigm to approach the new era with two dichotomous views apparent; (1) human-centeredness, and; (2) co-performance. These viewpoints are now described.

A large part of the developments in the history of the HCI field are geared towards designing for human-centeredness. [35] Human-centred design advocates to model users' natural behaviour into interface design so that it becomes intuitive, easier to learn, and with less performance errors.[36] Fundamentally, it is "an affirmation of human dignity…and an ongoing search for what can be done to support and strengthen the dignity of human beings as they act out their lives in varied social, economic, political, and cultural circumstances." [37] This view is widely adopted both in academia and practice. For instance, to apply these principles to designing solutions driven by AI, in July 2017, Google officially established their People in AI Research (PAIR) initiative, which aims to conduct, "fundamental research, invent new technology, and create frameworks for design in order to drive a humanistic approach to artificial intelligence".[38] Consequently, they regularly release projects and resources that help designers become acquainted with the possibilities AI offers.

Another prominent design company that announced their intention is IDEO. In late 2017 IDEO acquired the data science company Datascope with the ambition to, "create an offering we're calling D4AI: Design for Augmented Intelligence, which will be able to extend the capabilities of humans in a way that feels natural to them."[39] In addition, Microsoft is already applying their inclusive design principles to the development of AI.[40]

A recently introduced paradigm on the role design can play in the creation of solutions powered by ML is that of co-performance. Unlike the human-centredness view, the artefact in what Kuijer and Giaccardi[41] term as co-performance, is seen as "capable of learning and performing a social practice together with people." There is a direct link between decisions made during the design process and use practices carried out after. The locus of design is thus shifted towards solutions that allow for a recursive relation between design and use, allowing more room for evolving complementary capabilities and doings. Kuijer and Giaccardi argue that the concept of co-performance shows potential to be developed into a range of design approaches and tools that can aid designers of computational artefacts acknowledge that appropriate practice varies over situations and changes over time. The project *Resourceful Aging*, funded by the Netherlands Organisation for Scientistic Research demonstrates how ethnographic research paired with insights from co-performing household 'things' offers a way to design with ML. The project led to the design of data-enabled products, services propositions and simple interventions that promoted the vision; *ageing is an achievement and should be celebrated.*[42] The question still remains; how might these products, services and systems grow with users in surprising and delightful ways overtime; without diverging unintendedly toward harm.

Infrastructure as a necessary design material

35. Noessel, Christopher. "Designing Agentive technology: AI that works for people". *Rosenfeld Media*, 2017.
36. Oviatt, Sharon. "Human-centered design meets cognitive load theory: designing interfaces that help people think." In Proceedings of the 14th ACM international conference on Multimedia, pp. 871-880. ACM, 2006.
37. Buchanan, Richard. "Human dignity and human rights: Thoughts on the principles of human-centered design." *Design Issues 17*, no. 3 (2001): 35-39.
38 Wattenberg, Martin., and Fernanda Viégas. "PAIR: the People + AI Research Initiative," The Keyword (blog), Google, July 10 2017. https://www.blog.google/technology/ai/pair-people-ai-research-initiative/.
39 Budds, Diana. "Exclusive: Ideo's Plans to Stage an AI Revolution." *Fast Company*, 17 October, 2017.
40. Chou, Joyce., Oscar Murillo., and Roger Ibars. "How to Recognize Exclusion in AI." Medium (blog) Microsoft, September 26, 2017. https://medium.com/microsoft-design/how-to-recognize-exclusion-in-ai-ec2d6d89f850.
41. Kuijer, Lenneke., and Elisa Giaccardi. "Co-performance: Conceptualizing the role of artificial agency in the design of everyday life." In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, p. 125. ACM, 2018.
42. Nicenboim, Iohanna., Elisa Giaccardi., and Lenneke Kuijer. 2018. "Designing Connected Resources for Older People." In Proceedings of Designing Interactive Systems Conference 2018 - DIS '18. ACM Press.

We have introduced the notion of ML-powered solutions inadvertently transitioning into the area of wicked problems and the design as a discipline that can address this transition. However, choosing to view the problems ML addresses both deliberately or inadvertently as either tame or wicked means that an important aspect of AI's dual nature is overlooked. To be able to fully mitigate undesirable outcomes whether in the narrow state of AI or its possible incarnations, a holistic view must be adopted.

We turn to the construct of *infrastructure* as a means to simultaneously address tame and societal issues in AI. Prominently discussed by Star and Ruhleder,[43] infrastructure is a combination of interrelated social, technical and organizational arrangements.[44] As such it is composed of a complex matrix of objects and standards without absolute boundary or a priori definition.[45] The construct is familiar to the design discipline. An active area in the field of Participatory Design has been built around the notion of infrastructuring.[46] Bjögvinsson and colleagues define infrastructuring as an opportunity to extend design towards an open-ended, long-term, and continuous process.[47] The approach facilitates the emergence of new design opportunities by deliberately designing indeterminacy and incompleteness into solutions.[48] This leaves space for unanticipated events and performances so that solutions designed at project time can design boundary-objects (infrastructure) supportive of future design (at use time); essentially creating a chain of one design after another.

This notion of design after design[49] is well-suited to the nature of AI-powered solutions where users' regularly develop their own functionality, teaching an algorithm their preferences to create highly-personalised experiences. Choosing which songs to play, for how long and how often to listen to them, the user implicitly instructs a platform like Spotify to adapt to his/her needs; designing after it has been designed. This symbiosis between tame and wicked problems complemented by inadvertent designing-after-design can be made explicit by deliberately building infrastructure. To realise this proposition, we deconstruct the three main dimensions of infrastructure identified in literature and offer ways to build structures that can support the notion of design-after-design. These elements are: social, organizational and technological.

1. Social

The design discipline is already well-equipped to deal with this element through human-centeredness or co-performance perspective. However, there is still progress needed toward devising AI use cases that carefully consider human sensitives; both functional and ethical. The efforts of Microsoft, Google and IDEO are geared toward precisely this. A series of use cases by Microsoft Inclusive Design[50] exploring the requirements of an AI chatbot in various scenarios where children interact with the technology, discerned how design could be applied to identify and reduce bias – a practicality of design advocated by Jeanne Liedtka.[51] Joyce Chou and colleagues at Microsoft argue that design as an activity plays a critical role in developing an ethical framework based on human requirements for all potential users, not just a lead user

43. Star, Susan Leigh., and Karen Ruhleder. "Steps toward an ecology of infrastructure: Design and access for large information spaces." *Information Systems Research 7*, no. 1 (1996): 111-134.
44. Bødker, Susanne., Christian Dindler., and Ole Sejer Iversen. "Tying knots: Participatory infrastructuring at work." *Computer Supported Cooperative Work 26*, no. 1-2 (2017): 245-273.
45. Star, Susan Leigh. "Infrastructure and ethnographic practice: Working on the fringes." *Scandinavian Journal of Information Systems 14*, no. 2 (2002): 6.
46. Karasti, Helena. "Infrastructuring in participatory design." In Proceedings of the 13th Participatory Design Conference: Research Papers-Volume 1, pp. 141-150. ACM, 2014.
47. Bjögvinsson, Ehn., and Hillgren. "Design Things and Design Thinking." 2012.
48. Hillgren, Pers-Anders., Anna Seravalli., and Anders Emilson. "Prototyping and infrastructuring in design for social innovation." *CoDesign 7*, no.3-4 (2011): 169-183.
49. Bjögvinsson, Erling., Pelle Ehn., and Per-Anders Hillgren. "Design things and design thinking: Contemporary participatory design challenges." *Design Issues 28*, no. 3 (2012): 101-116.
50. Chou, Joyce. , Oscar Murillo, and Roger Ibars. "What the Kids' Game 'Telephone' Taught Microsoft About Biased Ai." *Co.Design*, October 12, 2017.
51. Liedtka, Jeanne. "Perspective: Linking Design Thinking with Innovation Outcomes through Cognitive Bias Reduction," *Journal of Product Innovation Management 32*, no. 6 (2015): 925-938.

or set of users.[52] Moreover, design encourages deeper consideration of the user's desires and emotions in context – with intention to build knowledge beyond preliminary user requirements. Where physical prototypes are not possible; Peter Lloyd also notes the power of imagination as a means to explore ethical considerations during the design process.[53] Infrastructuring may be a viable first step toward crafting technology use cases that explore social bias in order to mitigate unethical AI-development.

2. Organizational

To allow for design-after-design, a specific organizational structure that supports adaptability needs to be in place. The discipline has made strides, particularly using Design Thinking, in creating economic impact for businesses. In the context of AI, we contend a viable first step toward creating infrastructure that supports design-after-design to be the adoption of organizational ambidexterity's principles.[54] An increasingly popular construct for achieving long term firm survival,[55] organizational ambidexterity is defined as, "the ability to simultaneously pursue both incremental and discontinuous innovation. (…) Hosting multiple contradictory structures, processes, and cultures within the same firm."[56] Organizational ambidexterity allows companies to simultaneously manage current business demands and adapt to environmental changes.[57] We believe one form of ambidexterity in particular, design-led, to be well-suited for a context where a solution is never fully complete, rather in a state of continuous reconfiguration based on new insights.

An alternate funding model for infrastructuring design must also be considered. AI-powered solutions are continuously shape-shifting. The design firm must be attentive to the ethical and legal condition of the solution at any moment. Further, the design firm must be ready, both technically and financially, to intervene and correct an AI-powered solution. While the billable hour funding model is standard practice for design agencies and consultancies, a retainer model could prove to be a better fit for design-client management when dealing with AI. The retainer model allows a design agency to bill their client each annum (or quarterly) to 'retain' their services. The design agency can access this retainer funding to complete scheduled audits of AI-powered solutions on a continuous basis. Where the billable hour model is required again, perhaps for major updates or specific projects, it can be reintroduced as necessary.

3. Technological

To be able to address this part of the infrastructure, designers need to understand the technology (particularly its core – the algorithm) they are going to work with and be able to "converse" with it. Much like one needs to know programming languages such as HTML and CSS to be able to translate their design into a digital product (a website), designers need to understand the "new material" they're going to use. Steps are already being taken in that direction such as multiple emerging meetup groups between AI developers and designers around the world, companies and consultancies publishing their principles on how to design with AI (e.g. Google, IDEO, Microsoft, Fjord) and speculative design exhibitions to give shape to AI's possible incarnations (e.g. IDEO's exhibition Hyperhuman).

However, AI algorithms are oftentimes a mystery even to their own developers. This posits a question: how to understand something that is inherently opaque. We believe initial answers might be hidden in design cognition. At first sight, algorithmic logic appears furthermost detached from the discipline of design. Practically this may be so. Theoretically, formulating algorithms and the nature of design

---

52. Chou, Murillo., and Roger Ibars. "What the Kids' Game 'Telephone' Taught Microsoft." 2017.
53. Lloyd, Peter. "Ethical Imagination and Design." *Design Studies 20,* no. 2 (2018): 154-168.
54. O'Reilly III, Charles A., and Michael L. Tushman. "Organizational ambidexterity: Past, present, and future." *Academy of Management Perspectives* 27, no. 4 (2013): 324-338.
55. Oehmichen, Jana., Mariano L.M Heyden., Dimitrios Georgakakis., and Henk W. Volberda. "Boards of directors and organizational ambidexterity in knowledge-intensive firms." *The International Journal of Human Resource Management 28*, no. 2 (2017): 283-306.
56. Tushman, Michael L., and Charles A. O'Reilly III. "Ambidextrous organizations: Managing evolutionary and revolutionary change." *California Management Review 38*, no. 4 (1996): 8-29.
57. Mom, Tom J.M., Sebastian P.L Fourné., and Justin J.P Jansen. "Managers' work experience, ambidexterity, and performance: The contingency role of the work context." *Human Resource Management 54*, no. S1 (2015): s133-s153.

cognition share closer proximity. One of the fundamental paradigms of design methodology and consequently cognition – design as a rational problem-solving process – originated within the field of AI and was introduced to the design discipline by Newell and Simon in the early 1970s.[58,59] After this moment the two disciplines developed independently from each other. In a turn of events, the two disciplines meet again with AI-powered solutions entering wicked problem domains. We believe that principles developed in design cognition could be translated back to the devising of algorithms. However, this proposition requires the ability to overcome the problems with the widely criticised areas of Newell and Simon's paradigm, such as failing to account for the action-oriented, often implicit, knowledge associated with design.[60] Further research is necessitated to reach a deeper understanding of design cognition and its possible future incarnations and implications for AI..

Conclusion

This article has shed light on the deeper consequences of AI development for the design discipline. We have argued that to ensure narrow AI will have beneficial outcomes, design and its underlying rationale can create a new paradigm for development. We identify that there are fragments of methodological readiness for AI within the discipline; HCI, co-performance, thing-centred design and an attention to ethics. We have brought these methodological developments together around the notion of *infrastructure* in order to propose how to design with the new subject matter of AI. This is especially important if ML will be the most important way to improve user experience.

The strength and novelty of our proposition stems from the proposal to design infrastructures instead of solutions by understanding the interrelations and implications of these interlocking elements; social, organisation and technological. Further research is needed on their symbiosis and each separate element. We believe one way to do so is by continuously setting up small design experiments to prototyping infrastructures. The designer begins using current knowledge of prototyping (experiences, product service systems, organizations) to continuously test assumptions. The designer becomes responsible for shaping solutions beyond primary function; anticipating and evaluating new horizontal functions in collaboration society, organisations and technology. In development, engaging customers and stakeholders to define the problem and create shared value. In production, connecting manufacturers and material scientists in the physical realm, to programmers and system architects in the digital realm. Further, envisioning possible types of information that can be collected, and how the processing and sharing of that information might enrich the user's life well beyond a current set of lived and potential situations.

Thing-centred design and the notion of co-performance allows for new arrangements between users and objects that are AI-powered. What is promising is the collaborative efforts of industry and academia to engage in and share scientific breakthroughs regarding AI. Our proposition of harnessing infrastructure to approach AI is one step forward. Many more steps will be required to reach methodological clarity. What is clear, is that age of AI is here and that design will play a vital role in creating AI-powered solutions that grow with people, not against them.

58. Newell, Allen., and Herbert Alexander Simon. *Human Problem Solving.* (Englewood Cliffs, NJ, Prentice-Hall: 1972).
59. Dorst, Kees. "The problem of design problems." *Expertise in design* (2003): 135-147.
60. ibid