A Perspective on AI and Data in Design
Interview With Kate Crawford

Kate Crawford
USC Annesberg, Los Angeles
Microsoft Research, New York
kate.crawford@usc.edu

Vladan Joler
University of Novi Sad
Share Foundation
joler@labs.rs

Andrea Cattabriga
Alma Mater Studiorum
— Università di Bologna
andrea.cattabriga@unibo.it
ORCID 0000-0003-0975-2025

In this comprehensive interview, Professor Kate Crawford discusses the complex and pervasive impact of Artificial Intelligence (AI) systems on our global ecosystem, critically examining the material realities and extensive influence of these infrastructures. Delving into the socio-political and historical aspects of design in the context of AI, she shows that design is intrinsically linked to the consequences of capitalism, colonialism, and the concentration of power in technological systems. She challenges designers to acknowledge these connections and stimulate critical discussions about the role of design, promoting a vision where they can encourage diversity, challenge AI-driven homogenization, and question the processes of constructing meaning and decision-making in technology. Prof. Crawford posits that our relationship with technology is integral to our futures and underscores the importance of individual and collective politics in shaping these outcomes. The discussion advocates for a shift from technology-centrism to prioritizing collective planetary needs. It urges us to consider what kind of world we want to live in and what role technology should play in it.

Keywords
Artificial Intelligence
Ethics
Technology
Generative AI
Socio-technical systems

Professor Kate Crawford is a leading international scholar of the social implications of AI. She is a Research Professor at USC Annenberg in Los Angeles, a Senior Principal Researcher at MSR in New York, an Honorary Professor at the University of Sydney, and the inaugural Visiting Chair for AI and Justice at the École Normale Supérieure in Paris. Her latest book, *Atlas of AI* (Yale, 2021) won the Sally Hacker Prize from the Society for the History of Technology, the ASSI&T Best Information Science Book Award, and was named one of the best books in 2021 by *New Scientist* and the *Financial Times*. Over her twenty-year research career, she has also produced groundbreaking creative collaborations and visual investigations. Her project *Anatomy of an AI System* with Vladan Joler is in the permanent collection of the Museum of Modern Art in New York and the V&A in London and was awarded with the Design of the Year Award in 2019 and included in the Design of the Decades by the Design Museum of London. Her collaboration with the artist Trevor Paglen, *Excavating AI*, won the Ayrton Prize from the British Society for the History of Science. She has advised policy makers in the United Nations, the White House, and the European Parliament, and she currently leads the *Knowing Machines Project*, an international research collaboration that investigates the foundations of machine learning.

The purpose of the interview was to gather insights on the relationship between design and AI systems from a world-class expert who could provide the design community with some high-level thoughts and guidance on how this community should respond to the challenges it faces. The choice of a dialogue with Prof. Crawford stems from the need of the co-guest editors to talk about the deeper and less obvious aspects of this socio-technological phenomenon, and in particular about how the fundamental issue of data, which cannot be dealt with without touching on its technical materiality, brings us back to large-scale reflections in which disciplinary fences shift and call upon us to cross them. In this interview, she delves into a multitude of pressing issues that are highly relevant to the design field. These include the ethical ramifications of expansive training datasets, the legal complexities surrounding copyright in generative AI, and the sociopolitical implications of synthetic content. Given her extensive policy advisory roles, including engagements with the United Nations and the European Parliament, Crawford's insights are particularly timely. She advocates for a shift from technology-centric paradigms to ethical and collective frameworks, urging designers to grapple with the deeply embedded political and ethical dimensions of AI systems. The semi-structured interview was conducted in one remote session and was prepared by sharing questions in advance to make the discussion denser and more effective but leaving room for any thematic deviations. The structure faithfully reflects the discussion and the order of the interventions.

VJ For many years, one of your academic and artistic interests has been related to training datasets[1] and their impact on society. Even before the explosion of generative AI models[2], critically investigating datasets was an extremely difficult task. And now, in the past few years, we are witnessing exponential growth in the size and complexity of datasets. What new challenges are we facing in that field?

1
In AI systems, "training datasets" refer to collections of data used to teach machine learning models how to perform specific tasks, that can include text, images, or numbers. The model analyzes this data to identify patterns and make predictions or decisions, so the quality of the training dataset is crucial, as it influences the system's accuracy and potential biases. The implications of the use of data in design practice and research is relevant and is impacting a very broad spectrum of issues, such as the interpretation of data as material, design choices and human rights, platform-based service design, creative thinking and civic participation, to name just a few.

2
Generative AI is a family of technologies that uses machine learning algorithms to autonomously create new content, such as images, music, or text, based on patterns and examples from existing data. For an introduction on generative AI in relation to the design field see Thoring, Huettemann & Mueller (2023).

diid No. 80 — 2023
Doi: 10.30682/diid8023b

KC Training datasets form the foundation of how AI systems interpret the world. They set the epistemic boundaries, so I believe that they are important to investigate. Back in 2016, when I really began studying training datasets in detail, they were quite large, ranging anywhere from 10,000 items up to about 14 million in the case of ImageNet. But even then, you could do this very manual work of sifting through the later layers of data, cataloging their principles, and understanding the taxonomies. In a project I did with the artist Trevor Paglen[3], we called this work a type of "dataset archeology": excavating the data and sifting through it. It was incredibly time consuming, but very revealing, if you want to understand the underlying logics that an AI system has been trained on.

However, with the advent of generative AI, training datasets have now ballooned to an immense scale. For instance, one of the training datasets used for Stable Diffusion[4], a text-to-image model, is called *LAION-5B*, and it has 5 billion images and text captions scraped from the internet. How do you research something that has 5 billion images? This is almost reaching the limits of what can physically be achieved. In the case of ImageNet, if you spent 10 seconds looking at each image you could do this in five years, if you really wanted to, but in the case of *LAION-5B*, it would take you almost 2000 years.

In some ways, from a manual perspective, we're reaching an event horizon of datasets where they've become so massive, they are like enormous gravitational objects that absorb all light and all investigative possibilities. This is a challenge to investigators, but it's one that we can meet. And I remain optimistic about this work, and its importance, because we can begin to develop our own tools, our own technical systems for investigation. In our international research team, *Knowing Machines*[5], we've created a technical lab and we've been developing tools to study datasets. That work is led by Christo Buschek who recently was awarded a Pulitzer Prize for his work on data investigations. We're releasing those tools as open-source toolkits, so that more people can join us in this process of studying datasets.

VJ One of the many shady aspects of contemporary generative AI models' production is certainly the issue of copyrights. The training of LLM[6] relies on a huge amount of data that is being massively collected through non transparent processes. We are talking about tens of thousands of books, academic articles lines of code and the complete history of art, which are now being extracted and used without much questioning. We saw similar practices much earlier, when Flickr images under the Creative Commons licenses were scraped[7] to train face recognition models. It looks like both the copyright and copyright side of the spectrum are failing to keep up with the present situation. What is your view on this issue?

KC It's interesting that many debates about generative AI have centered on the crisis of creative labor, or perhaps the crisis of humanity itself. However, the more immediate crisis lies in

3
The reference is to the 2019 project *Excavating AI*, consisting of an essay (Crawford & Paglen, 2019b) of the same name and an exhibition at Fondazione Prada in Milan.

4
Stable Diffusion is a text-to-image model capable of generating photo-realistic images given any text input (https://en.wikipedia.org/wiki/Stable_Diffusion).

5
Knowing Machines is a research project sponsored by the Alfred P. Sloan Foundation, investigating the training of machine learning systems and their impact on interpreting the world. The project aims to develop critical methodologies, tools, and communities focused on understanding training datasets and their role in constructing "ground truth" for machine learning (https://knowingmachines.org/).

6
Large Language Models (LLMs) are a specialized type of artificial intelligence model in the field of natural language processing, based on statistical methods, that construct text based on probabilities derived from extensive training data.

7
Data scraping refers to the automated extraction of images from websites, often using specialized software. While efficient for gathering large datasets, this practice raises ethical and legal concerns, including issues of copyright and user consent.

copyright law. Because for the last 300 years, Anglo-American copyright law has sought to use copyright as an incentive for people to produce work. But now we have a very strange situation, where so much of the work that's produced by generative AI is training on the work of centuries of artists and writers, and creative producers of all stripes. Currently, both the US and EU courts consider the use of such data as fair use. However, this is continuously being challenged. For instance, this week, a massive lawsuit was launched against OpenAI for allegedly using copyrighted books to train ChatGPT[8]. While it's uncertain if the lawsuit will succeed, it's pushing the boundaries of what copyright law can do. Certainly, fair use[9] itself was not designed to allow large tech companies to harvest the entire internet, this ongoing capture of the commons. It was designed for artists to be able to build on prior work and to make transformative changes. So, fair use has indeed been stretched to its limit. But there's something even weirder going on with the outputs of generative AI. According to US law, non-human authors can't have copyright. There was a well-known case where a monkey took a selfie by pressing a button on a photographer's camera, and it was determined that the monkey did not own this image, because nonhumans cannot hold copyright (Domonoske, 2016).
This is also applied now to generative AI. Consequently, everything produced by these systems such as Stable Diffusion or Dall-E has no copyright, even when prompted or given instructions by humans. This is actually quite radical, because it means that right now, billions of artworks are being produced without any recognizable author. We could think of this as the largest global experiment in art *after* copyright. That's a pretty extraordinary change. And all of that work is now unowned and can be used anywhere by anyone for anything. So, in this sense, we have a moment of profound change. It's taken me back to Walter Benjamin in 1935 (Benjamin, 1969), when he was facing what he saw as an Age of Mechanical Reproduction. He asked what this meant for the aura of the work of art, and for him, it meant that, essentially, the work of art had already become detached from the domain of tradition. And now what we're seeing is that the work of art is not just detached from tradition, it's even detached from human artistic production of any kind. It's going into this very weird space, and I think we are seeing a crisis of what copyright is for and how human artistic production circulates.

VJ Indeed, but also on multiple different levels, creatives spend years creating something, only to be told that a statistical projection of their work is not considered their work. That's a complex issue!

KC Copyright law's three central concerns — authorship, expressiveness, and agency — are all facing enormous epistemological challenges from generative AI systems. These issues can only be resolved politically, not by tech companies or artists. Politicians will have to decide how or whether to reg-

8 ChatGPT is a language model developed by OpenAI for generating human-like responses in conversational settings (see https://openai.com/blog/chatgpt).

9 "Fair use" is a legal doctrine in U.S. copyright law that allows limited use of copyrighted material without requiring permission from the rights holders (Fair Use, 2023).

diid No. 80 — 2023
Doi: 10.30682/diid8023b

ulate these systems and who will benefit from generative AI. And right now, the greatest benefits are going to the smallest number of people in Silicon Valley, and are not being shared with the artists, creators and programmers whose work shapes these systems.

VJ As we anticipate a probable flood of synthetic content in coming years, this automated hyper-production can create numerous problems that we cannot even predict. It's easy to imagine countless dystopian scenarios. Over the last few years, fake news has been a hot topic in academic and regulatory circles. With the rise of generative AI, we have essentially perfected tools for the automation of fake news production, eliminating the need for armies of trolls. What scenarios do you envision and what are your thoughts?

KC First of all, the synthetic loop problem where AI systems are trained on AI-generated content is already happening. We're already starting to see a flood of synthetic content online, and it means that we could be reaching a type of law of diminishing returns in terms of how effective generative AI can be. Multiple studies have shown that if you start to mix synthetic data — generated content —, the models start to fail, they get very skewed very quickly. This could potentially create a collapse of meaning as machines fail to detect what is generated and what is not. Their own systems of meaning-making could start to crumble. So, I think there is something really interesting here about the precipice on which we stand, because we are about to move from a time before when the internet was primarily content generated by humans, to a moment where most of the content online will be generated. I think that's the line that you can draw right now, and I agree that there's a lot here that we cannot predict, but what we can certainly know is about to get a lot weirder.

The knowledge economy of fake information has now permanently changed. You know, truth and reality have always been complicated historically, for centuries, but now they are collapsing epistemically, and I think the full implications of that are profound. I haven't seen anyone in the tech sector really contend with what that means.

More obvious concerns include the impact on information during an election and our news ecosystems. For instance, when people are generating a perfect replica of the *New York Times* but with completely different stories, and you won't be able to tell the difference. However, there's something even bigger at stake, which is the atomization of the reality bubble down to such a small level that there is no way to even start to fact check against what becomes the reliable truth. I think we're losing the idea of a stable point of reference. So even bigger than the sort of acute crisis around elections and the fourth estate, we have to think fundamentally about how we perceive the world, how we perceive each other, and that things will rapidly become very unstable.

VJ Ironically, the same companies creating these tools for the mass production of misinformation might also be the ones we look to for solutions. They are stuck in a strange loop, simultaneously creating the problem and being tasked with finding its cure.

KC Yes. Do you remember the animated film[10] the *Sorcerer's Apprentice*, where Mickey Mouse is the apprentice? He tries to create a magic spell so brooms will do his household chores, but this results in more problems — the bewitched brooms go out of control. I think we're much closer to that scenario, because the technology companies that are attempting to find a way to create shortcuts are creating new, wicked problems. I just keep thinking about that moment when Mickey Mouse has the realization and you can see his face freeze, and it's this moment of "what have I done?". That moment is here, because there isn't a quick fix or a magic button to differentiate between real and generated content. Watermarking technology or reliable assessment tools that can determine whether content is real or AI-generated, like ChatGPT, are currently not reliable. Therefore, no one is coming to save us and the automated spell-casting attempts by technology companies are going to create new problems that we can't always predict or see coming.

AC In 2018, together with Vladan you published a large-scale map and the accompanying paper titled Anatomy of an AI System (2018). A big part of that map and the story you explored back then was related to the materiality of AI systems and the vast frontiers of a planetary scale of extraction. Five years later, what main differences or changes are you seeing in those systems? And how can designers contribute to the creation of a more sustainable and ethical AI system and infrastructure from the level of the product service ecosystem?

KC Since we produced that map, all the mechanisms we studied have become more intense. For example, recent research shows that generative AI is at least five times more energy intensive than traditional search. One exchange conversation with ChatGPT is the equivalent of pouring a bottle of fresh water into the ground. This is terrifying at a time of extreme climate crisis.

For us that project felt prophetic in that it captured AI's inherent dynamics as being deeply connected to supply chains, logistical processes, environmental and human labor extraction processes, which have expanded exponentially since we produced that work five years ago. What's been so interesting about that project is its long life. It's still being displayed in museums and galleries around the world; it opened at MoMA in 2022 and it will stay up until 2024. Its relevance has grown due to generative AI pushing these dynamics to this extraordinary point of acceleration that we're all experiencing.

To answer the second part of your question, what can designers do — and I'd be curious to hear what you would say Vladan — I wouldn't call myself a designer, but I would call myself somebody who's deeply invested in the politics of design. And what that means is that I think designers have an

10
Disney, W. (Producer), & Algar, J. (Director). (1940). *The Sorcerer's Apprentice* [Film segment]. In Walt Disney (Producer), Fantasia. Walt Disney Productions.

diid No. 80 — 2023
Doi: 10.30682/diid8023b

enormous role to play in revealing the systems underneath the sort of shiny, smooth surfaces of the technologies that we use every day. But also pushing back and saying we don't have to use generative AI for everything: it's like using a battleship where you could use a pair of scissors, it just doesn't make sense. And I think designers are very good at thinking about minimalism, thinking about using what is needed, where it is needed, rather than grotesque examples of a type of hyper indulgence and lack of concern for environments and populations. If we think of all the ways that we use ChatGPT, in many cases they are actually not what it's designed for, not even what it's good at. For example, it's not a good fact generator or a search engine. There are some things for which people are simply using it as a gimmick or as a trick, and it's a very expensive magic trick to be using in that way. But I'm curious, what do you think, Vladan?

VJ Design is a really powerful tool and power that we can use to rethink things differently, to rethink technology, to think more sustainably about what we are producing. But I think it could also become a kind of limited perspective, in the sense that it's not just a "design problem". Yes, it's a much wider problem that we need to solve, so we need to rethink the values. But the problem that we addressed with *Anatomy of an AI System*, as well as with *Knowing Machines*, is a kind of deep genealogy or history of wrong values. It's naive to think we can solve the world's problems with design if we don't address issues such as inequality and exploitation that are embedded in how these technologies are made. Design-oriented projects like *Fairphone*[11] have tried to tackle these issues, but they struggle to penetrate more than a few steps into the supply chains. Yet, we must also acknowledge the existence of darker patterns in design as well. It's not just the interface that determines what we see or don't see. The influence of design is incredibly profound. Moreover, in most instances, it carries significant political implications.

KC It's always political and this is what I love about the work that we've been doing together. But while *Anatomy of an AI System* looked across space — mapping the supply chain components of an AI system, from mines to server farms, to exploited workers — this new project called *Calculating Empires*, looks across time. By looking at five centuries in the relationship between technology and power, we can ask really different questions about how change is possible. It's no longer design that can fix all problems, because design is embedded within the economic systems of capitalism, and within historical processes of colonialism. It is embedded in these technologies that are themselves implicated in an accelerating centralization of power. So, in some ways, I think that my answer makes these questions more complex, it makes them more honest. You can't give easy quick answers about how design solves problems without confronting its history, without confronting its indebtedness and without confronting its enmeshment in larger economic practices. And that, for me, is a much more powerful conversation. It's like we're having a real talk about what design can and can't do. So this project we are working on now, I think, has really shifted my thinking, in a truly challenging but profound way.

11
Fairphone is a sustainable smartphone initiative promoting ethical manufacturing, responsible sourcing, and extended product lifespan (https://www.fairphone.com).

AC Indeed, with Anatomy of an AI System, you've shed light on the distribution and even unintended consequences of AI systems. They can displace effects from the point of interaction to exploited workers in distant regions such as Kenya or South-Eastern Asia. Perhaps it will be important to consider this displacement of effects not only in the present, across places, but over the coming years. As you know, there's a relatively new framework called Responsible Research and Innovation, backed by the EU Commission, which is becoming the reference for all European policy initiatives related to research and innovation. It's not just about avoiding issues but about embracing responsibility. In this context, how can we invoke professional responsibility as well, and how can policies shift towards a more responsible approach? I believe this will be a crucial aspect to address.

KC I agree. One of the things that I write about in the *Atlas of AI* (Crawford, 2021) is that so much weight has been put on the concept of ethics, but actually, while ethics is necessary, it is not sufficient. We have to focus on the issue of power. How is power distributed? Who has the power to change things? Who does not? Who benefits from systems and who is most exposed to risks and harms? The minute we shift to a power analysis of these systems and their social impacts, we get very different answers than if we say "oh, how do we make sure we design this algorithm ethically?". A sole focus on ethics tends to narrow our approach. It was indeed a controversial statement in the book and irked many practitioners in AI ethics. But I argue that if you're not addressing power, you're missing out on the dynamics of what's really going on. This is precisely why I'm currently engaged in a project with Vladan, heavily focused on history, as it allows us to trace the threads of power over time.

AC As AI offers the potential to bridge gaps between science, folk wisdom, and varied ways of knowing and perceiving the world, there's a risk that we may simplify human existence to make it easier for artificial agents to interpret our world. This could lead to imposed homogenization under the guise of personalization of micro-experiences. In this context, what advice would you give to designers who are shaping AI-driven processes of meaning-making or decision-support to ensure diversity and prevent homogenization?

KC One of the concerns that I have focused on for many years now — since doing the *Training Humans* show with Trevor Paglen in 2019 (Crawford & Paglen, 2019a) — is the way in which AI systems are not just personalizing our experiences and making things easier, adapting to us, but the fact that we are adapting to them. Humans are the ones being trained to fit their thoughts into a ChatGPT textbox, not the other way around. We are having to learn the prompts and be prompt jockeys and adapt to creating desired images from Stable Diffusion rather than the systems adapting to our needs. So really looking at the ways in which, rather than being examples of machine learning they're examples of us learning to be more machine-like, that has been one of the central concerns of my work for a very long time.

diid No. 80 — 2023
Doi: 10.30682/diid8023b

Hilariously, one of the sorts of things I was thinking about was just how long I have been interested in this question. In my early twenties I was in an electronic band called B(if)tek, where we wrote a track named *Machines Can Do The Work*. It included a sample from a 1960s IBM film with the line, *Machines can do the work so people have time to think*. By the end of the track, we reversed the saying to "Humans can do the work so machines have time to think", reflecting how humans are providing all this labour to make machines appear as though they're thinking, a scenario reminiscent of the current state of generative AI.

So what guidance would I offer to designers contemplating these issues? The first step is to recognize the degree of homogenization occurring behind the scenes. Systems like Stable Diffusion and GPT-3[1] are trained on largely Western, heavily Americanized content. GPT-3 communicates like an American, while Stable Diffusion generates images based on a universe of averages. If we think of these machines as a universe of infinite averages, that's what they can offer us. What do you lose with that type of averaging? What sorts of stories can it not tell? What sorts of images can it not produce? This is where designers play a really powerful role in bringing that awareness back in. Generative AI cannot be the one single interface to truth that we simply accept, and just play with prompts around the margins. I think designers have really good ways of showing what we're not seeing with generic tools. And the importance of that is going to grow in the next few years. How can you do things the machines can't do? That is the question that's being posed to every designer right now.

AC As artificial intelligence technologies prove to be effective tools to address some of the contemporary pressure towards productivity and efficiency, designers might need to engage with the political impact of their proposals in order to contribute to cultural solutions, and contribute to radical approaches to social dilemmas. In this context, fifty-year-old insights by Tomás Maldonado on the interplay between design and technology seem to be extremely topical. He warned about the risks of assigning a magical role to technology without grounding it in social values to avoid mere solutionism, and pointed out the contradiction when design is caught between mature technologies and immature decision-making power centers in society. Today, as digital technologies rapidly shape social transformations, political structures struggle to keep up. Are we destined to become increasingly political in our individual actions to save humanity from a chain of unpredictable consequences, or will policies suffice to limit risks and unlock the positive value of these technologies?

KC I'd like to think that we will realize the importance not just of becoming individually political and living our values, but the essential practice of collective politics. Because I do not think that on the basis of what we have seen so far, governments are going to be able to keep up with the speed of generative AI systems and the cascading social implications. We are going to be faced with a profound reorientation of

12
GPT-3 (Generative Pre-trained Transformer 3) is a language model developed by OpenAI, serving as the foundation for ChatGPT, an AI-powered chatbot capable of engaging in natural language conversations.

13
Maldonado's unsystematized thinking about technology can be found in *Reale e virtuale* (1992/2015), *Critica della ragione informatica* (1997/2006) and *Memoria e conoscenza* (2005).

responsibility. How can we start to think collectively about where we will and will not use these systems and what risks we will and will not accept? What type of worlds do we want to live in? And to answer those questions first, we must think about what we want the world to look like, rather than thinking about what technology can do, the most important thing now. We need to decenter technology, rather than allowing it to drive our visions of the future.

References

Benjamin, W. (1969). The Work of Art in the Age of Mechanical Reproduction. In H. Arendt (Ed.), *Illuminations* (translated by Harry Zohn, from the 1935 essay). Schocken Books. http://web.mit.edu/allanmc/www/benjamin.pdf

Crawford, K. (2021). *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press.

Crawford, K., & Joler, V. (2018). Anatomy of an AI System. *Virtual Creativity*, *9*(1), 117-120. https://doi.org/10.1386/vcr_00008_7

Crawford, K., & Paglen, T. (2019a). *TRAINING HUMANS* [Exhibition]. https://www.fondazioneprada.org/project/training-humans/

Crawford, K., & Paglen, T. (2019b, September 19). *Excavating AI*. https://excavating.ai

Domonoske, C. (2016, January 7). Monkey Can't Own Copyright To His Selfie, Federal Judge Says. *NPR*. https://www.npr.org/sections/thetwo-way/2016/01/07/462245189/federal-judge-says-monkey-cant-own-copyright-to-his-selfie

Fair use. (2023). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Fair_use&oldid=1165943320

Maldonado, T. (2005). *Memoria e conoscenza. Sulle sorti del sapere nella prospettiva digitale*. Feltrinelli.

Maldonado, T. (2006). *Critica della ragione informatica*. Feltrinelli.

Maldonado, T. (2015). *Reale e virtuale*. Feltrinelli.

Thoring, K., Huettemann, S., & Mueller, R. M. (2023). The Augmented Designer: A Research Agenda for Generative AI-Enabled Design. *Proceedings of the Design Society*, *3*, 3345-3354. https://doi.org/10.1017/pds.2023.335