AI as a Design Material: Dealing with New Agencies

RAMON SANGÜESA
Technical University of Catalonia

ARIEL GUERSENZVAIG
Elisava School of Design and Engineering

KEYWORDS
Artificial Intelligence, Design, Materials, Agency, Machine Learning, Ethics.

HOW TO CITE
Sangüesa, Ramon and Ariel Guersenzvaig. 2019. "AI as a Design Material: Dealing with New Agencies". *Temes de Disseny* 35: 6-25.

1 INTRODUCTION

Artificial Intelligence (AI) has become a household word, a media buzzword but also an industry in itself, and a set of practices linked to business and professional goals. As a general symbolic technology with enormous investments behind it, it has impacted many spheres of activity. Chatbots populate the internet and are put to use for marketing (Wilson 2017) and political purposes (Bond et al. 2012), intelligent systems orchestrate logistic chains (Kückelhaus and Chung 2018), usher the attention of people towards products (Marinchak et al. 2018), influence their next purchase (Portugal et al. 2017), change their emotions (Booth 2014), interpret medical tests (Koch et al. 2018), interact with users in new ways (Siddike et al. 2018), decide who gets a grant or a loan (Agarwal 2019; Eubanks 2018), or when and how to fight war (Davis 2019). In some way or another, intelligent systems classify, schedule, plan, and build products, services, and experiences. They continuously make decisions regarding many aspects of our life, contributing to shaping it (Turkle 2006).

However, this is not the first time that AI techniques have undergone an upsurge in their application in industry. The 1980s and 90s saw a wealth of applications and high hopes for the technology but they ended in the first "Winter of AI" (AI Expert Newsletter; Russell and Norvig 2003), given the fragility shown in real use by the AI technologies of the time. Nowadays, the degree of digitization seen across entire sectors of activity and the improvement of many AI techniques have created a fertile ground for the widespread use of AI once again. Where the earlier versions of AI were driven by the so-called expert systems that were costly to develop and difficult to train, the combination of the current volume and availability of data for many types of applications has facilitated the use of Machine Learning techniques, historically a subdiscipline of AI. Machine Learning has sped up the construction and training of AI systems, overcoming the previous scarcity of sources from which knowledge could be extracted, and with which AI systems could perform their tasks.

There is a feeling nowadays that AI will play a central role in our societies. It is portrayed as a strategic asset that, as just one example, will impact international competition (Villani 2018; Webster et al. 2017; NSTC 2016; Hogarth 2018). AI has been tasked with an enormous influence, for better or for worse, in organizing everything from personal behaviors to economic strategies. Its connections with behavioral economics (Pedersen 2018), digital surveillance (Zuboff 2015), and military applications (Krishnan 2009) cast a dark shadow on the preeminence of AI as an orchestrator of life and have spawned intense activity concerning its ethical dimensions (Floridi 2019; Dignum 2018; Casacuberta and Guersenzvaig 2018).

Design as a professional practice and a sector in itself has also undergone the friction and the excitement of AI. Professionals in different areas of design have found themselves dealing with AI in several ways. Perhaps the first ones who felt the impact of AI were the design professionals involved in Human-Computer Interaction (HCI) and User Experience (UX) (Winograd 2006; Grundin 2009), historically the place where the digital impacted design most intensely. But AI techniques, and Machine Learning techniques in particular, have also made inroads in others aspects of design.

Many generative approaches to design make extensive use of the interconnection between data and algorithms, especially machine learning algorithms, to devise and produce new products in such a way that the intelligent algorithm has taken a role which goes beyond that of a passive or reactive tool in the hands of designers (Koch 2017). Also, design has had to tackle new problems and goals when dealing with recommender services and intelligent assistants. Slowly, we realized too that some designs are actually being developed not by or with Artificial Intelligence but for these types of entities (DiSalvo 2011). AI systems, then, sometimes simultaneously play the role of a tool, a collaborator, and a user.

We felt that this situation invited some reflection from the standpoint of design and design engineering, and we sought to find a common theme that could be useful to explore the relationship between AI and Design as well as other emerging technologies such as the use of Big Data or intelligent materials. AI is mainly understood as a "soft" entity but it also has a material component, similar to other areas of Information and Computing (Dourish 2017, 33-59). We thought that this could be a good starting point for contrasting different interactions between AI and Design and, also, to start a possible clarification of the current state of these connections. However, in doing so we also wanted to stress that as a design material, AI has very peculiar characteristics. It is difficult to understand AI as a material that is passive and constant. On the contrary, it is not only malleable but it has a certain degree of "life," of autonomy at least, by itself. Designing with AI is definitively working with a material that will evolve in ways that are more varied and more difficult to plan ahead, at least in comparison with other more traditional materials. Working with plywood or aluminum is fully bounded by the laws of physics, which makes them highly predictable materials. AI algorithms are not outside the realm of physics, of course, but their behavior, by their very nature, could be very difficult to anticipate in detail. There is, at the moment, no straightforward way to look into an AI model and make it fully predictable.

Is there a common concept that could help us consider, under the same light, the diverse types of design that are being done in connection with AI, Machine Learning, and data? Is there a concept used by both disciplines that could help us map the relationships between them? After some thought and deliberation, we concluded that the notion of "agency" was a good candidate to help us conceive a practical connection with the objects of AI, ML, Big Data, and Design. If nothing else, certain similarities in their conceptualization and use in practice in both disciplines were intriguing.

The difference that AI may introduce in general – but to design in particular – could be understood as a leap in the level of agency that the results from using AI as a material in design. In design, "agency" has been introduced and used in several subdisciplines and practices, for example, in design research but also to guide design practice. The notion of "agency" carries a certain tradition, for example, from fields such as Actor-Network Theory (Latour 2005; Yaneva 2009). In design, the agency of objects is recognized and may be connected to concepts such us affordances or might be applied in user research. In a way, agency is a concept accepted within design practice. Some concepts such as the idea of "Objects with intent" (Rozendaal 2016), for example, have strong similarities to the types of agency used in AI. One could say that AI is, in effect, now the discipline of creating artificial agents with their own goals, desires, and intentions. Scholars from the postphenomenological approach (e.g. Verbeek 2011) articulate a philosophical perspective in which (moral) agency becomes a matter of human-technology mediations rather than a strictly human issue.

Indeed, in Artificial Intelligence, the concept of agency and agencies has become central and is a basic construct to guide the construction of complex AI systems (Wooldridge 2002).

There are differences between the conceptualizations used in each discipline. In order to clarify them we could go back a little and explore the origins and goals of AI and why and how the concept of agency now plays a crucial role across disciplines.

2
AGENCY IN ARTIFICIAL INTELLIGENCE SYSTEMS

The beginning of Artificial Intelligence as a discipline is usually associated with the celebration of the Dartmouth Summer School on Artificial Intelligence in 1956 (McCarthy et al. 1955). Several researchers gathered there, representing mathematics, information theory, logic, computing, and psychology. They shared a common excitement about what computers could do. The Dartmouth pioneers envisioned these machines beyond their scope of use at the time. They thought that computers could be entities capable of doing much more than just crunching numbers. They thought that computers could manipulate many different types of representations, including representations of knowledge. Thinking could become, as Hobbes had said centuries earlier, just a matter of calculation. Although this view of calculation of new types of symbols, not just numbers, was the predominant one in the idea of machine intelligence that dominated the Dartmouth meeting, another part of the same original group was more interested in simpler calculations. Instead of dealing with complex symbolic structures they thought that replicating the simple signal combination of the neural infrastructure of intelligent, cognitive systems would lead to intelligence. Together, the Dartmouth group issued a declaration that, more than a precise program of scientific research, looked more like a list of examples of tasks that they associated with intelligent behavior and that they thought that should be explored. It is worth having a look at this list (McCarthy et al. 1955) since it sheds light on the ambition of the Dartmouth group vision but also on the limits of the imagination of that founding group: given the abilities of computers at the time, one had to be brave to

The Cyborg Hand, by Judit Parés. Elisava Final Degree project; Simultaneous Studies Program in Design and Industrial Design Engineering, 2018.

extract from them the extremely more complex tasks that making them intelligent could require in the future. It was an exercise in speculation and projection, perhaps inspired by a research milieu under the spell of the achievements that Alan Turing and others achieved in the 30s and that led Turing to write his seminal paper "Computing Machinery and Intelligence" (Turing 1950). The overall goal of the budding Artificial Intelligence discipline was summarized by one of the participants at this historical meeting as:

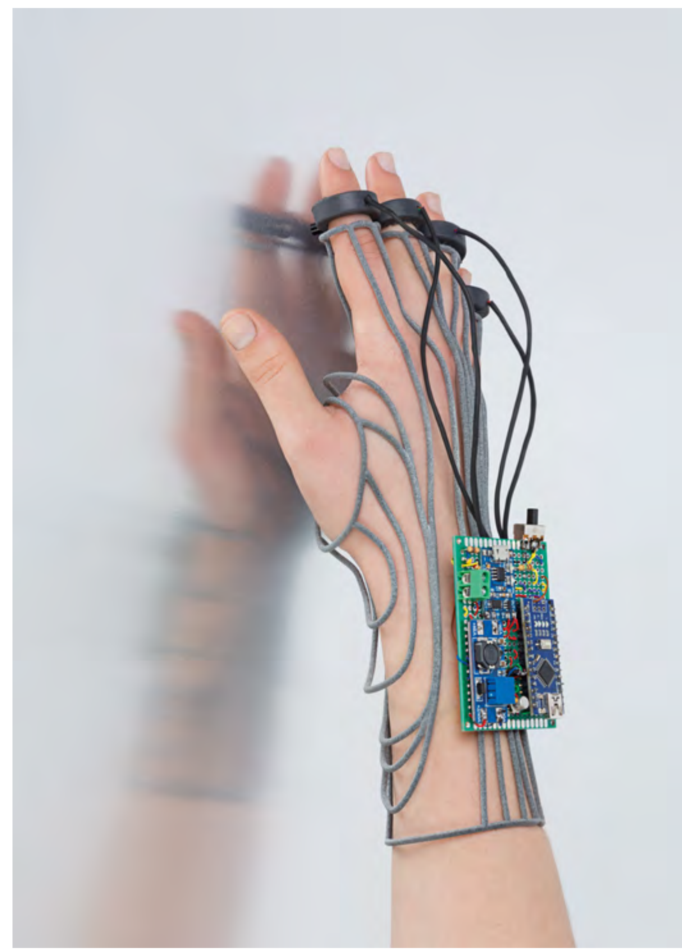
"*To study how to create systems that behave in a way that if they were humans it would be seen as 'intelligent.'*" (Minsky 1986)

Interestingly enough, as practitioners of a newly born discipline, their ways of working towards this goal departed from the practices of the disciplines each of the participants came from. More than a scientific methodology – understood as the methods of the theoretical or experimental hard sciences – they preferred a constructive approach or, in later terminology, a constructionist approach. That is, they would learn how those goals could be reached by building systems to attain them and learning from the process. These are essentially the traits of a "culture of design" (Serra 1992), as Herbert Simon (one of the Dartmouth meeting's most prominent attendees) would agree (Simon 1961).

After that meeting, a lot of activities started and many initial successes appeared rapidly. By the end of the 1980s and the beginning of the 1990s the most promising area

in AI was Expert Systems. They would become Knowledge Systems and gave birth to a whole new discipline, Knowledge Engineering. Expert Systems were essentially reasoning systems that used a representation of the knowledge accumulated through practice by very proficient human experts. This knowledge was very restricted to an area of practice and was connected to the performance of an intelligent task: diagnosis, scheduling, or design, for example. There was a lot of interest in business about these Expert Systems, from Insurance to Aerospace, and huge amounts of money were invested in their construction and maintenance. But the systems at the time exhibited brittleness and rigidity. They had enormous difficulties in learning new things. Consequently, they were not able to adapt to changes in their environment rapidly enough, sometimes failing abruptly. The field of AI was under attack (Lighthill 1973). Research funds were cut and the different practical solutions and methods devised seemed unconnected and conceptually disparate. The discipline was in total disarray. There was a profound questioning of its initial goals and assumptions. Some wondered if it was necessary to reason to perform intelligently or if reasoning needed a symbolic representation (Brooks 1991). Was logic actually needed to reason and think or would simply imitating the connectivity of neural tissues be enough to perform intelligent tasks? Incidentally, what was an intelligent task? What was the common concept between, say, natural language processing and robotics? Was there any?

It took a while for the discipline to reorient itself. The creation of a common concept to characterize its object of interest was instrumental in its slow rebirth. The concept



of Agent reorganized the interests and goals of the different subdisciplines (Russell and Norvig 2003). AI revolved around building intelligent agents that solved problems alone or collectively. In order to do so, agents had goals and applied any knowledge they already had or that they could gather from the environment they operated in. Agents had "agency."

There were two main understandings of this agency. One was an ability subservient to another agent, which could be a human agent. That is, an intelligent agent acted on behalf of another agent (possibly human). To do so, it had a wide range of alternative ways to reach its goals on behalf of this other agent (i.e. the goals of the first agent were a translation of the ones of the second agent). In a stronger definition, "agency" meant the autonomy of an agent to pursue its own goals. Agents could "come to life" whenever their environment showed some configuration, some state that would prompt them into action. They would do whatever they could (including learning) to attain that goal in that configuration of the world. Agents could be completely software- or hardware-based, most frequently a mix of both. Multiagent systems were collective intelligence expansions of the agent paradigm (Russel 2003). In any case, an agent in modern AI is a special type of rational agent (Simon 1961) that operates in an environment and pursues its goals. The agent knows about the state of the environment through sensors and operates in it by using effectors. Action is the result of decision making that is arrived at by mobilizing the set of beliefs of the agent that connect the known information of the environment with knowledge. In some definitions of intelligent agents, desires represent the situation or situations that the agent would like to attain to arrive at its goals. Intentions represent the deliberative state of the agent, what the agent has chosen to do.

The agent concept put every subdiscipline in AI in a clearer position by locating itself depending on its view of the agent concept. Therefore, one could understand robotics as the creation of hardware agents that should interact physically with a physical environment. Analogously, one could create a software agent that could operate in a virtual environment. The discipline had finally found a common way to portray and compare the different types of AI systems: the quality of their agency.

Agents, data and learning

One of the most important changes in the construction of systems using intelligent agents since this refoundation of Artificial Intelligence is the importance that machine learning methods have acquired. Increasingly, intelligent systems are built as a result of a process that starts with data that seem to be relevant to a given task. The data are used as the raw material for a machine learning system that extracts knowledge that can be implemented. For example, machine learning systems can help identify relevant patterns in an application domain. A machine learning algorithm is used to create a model that, when the system is later deployed, will let it decide if a given stream of information is representative or not of this pattern. This is the basis of classification systems that are routinely used in many applications, subsystems, and fully-fledged systems. For example, in identifying faces, classifying situations, objects, people, etc. In general, machine learning methods could be roughly divided into methods that

EIA: Espacio Interactivo Artesano, by Marian Brea, Nil Muriscot, Omar Josep Paul, and Nicolas Tarragó. Elisava Degree in Industrial Design Engineering, 2017.

give a criterion for clustering things together in unknown domains, classifying objects in domains where we know which are the main classes existing there, and, finally, methods that help an agent learn by itself how to improve its own performance, that is, getting better at a given task. Typically, these three types of learning are referred to as unsupervised learning, supervised learning, and reinforcement learning (Alpaydin 2017). There are other methods that do not fit so neatly in this classification, such as learning by analogy, metaphorical learning, or transferable learning. Another machine learning classification system would divide machine learning methods into analytical-descriptive or generative. That is, methods that help create a model to understand a domain and extract decision rules for the intelligent system to be deployed later based on that model, or methods that learn from data and create new things such as Generative Adversarial Networks, or GANS (Goodfellow et al. 2014).

Let us pause here to introduce an example to briefly explain the main differences between the earlier AI systems (expert- or knowledge-based) and the newer machine learning systems (based on big data and pattern recognition). Imagine you wanted to develop a system that could recognize a dog in a picture. The earlier systems would be trained to recognize characteristics of a dog (basic shape, ears, eyes, fur, nose, etc.) that were somehow defined beforehand. We could say, a dog has between zero and four legs. It has between zero and two ears, and the ears can be tiny or huge. To make a long story short, this never quite worked. The complexity of the world is simply too great to be described in advance. The new systems work the other way around, you train the system not by teaching what a dog looks like through abstractions, but by feeding it millions of pictures of dogs codified in a string of numbers. In the best case scenario, the system then learns what a dog looks like by recognizing numerical patterns that are present in those pictures. When you show the system a new picture of an animal, then the system can determine with a certain degree of accuracy whether that picture features a dog or not. It is important to note, that current ML systems, even if able to identify a dog, never really has a semantic understanding of what a dog is.

Within this landscape of methods and techniques, models based on neural networks have seen increased practical success over the last fifteen years and their application has become widespread. They are the basis of what is known as "Deep Learning Methods" (Goodfellow et al. 2016).

At present, most AI systems currently being deployed make very heavy use of machine learning techniques. That is, current AI agents tend to be learning agents with a higher possibility of autonomy than previous AI systems since machine learning can be triggered from new data whenever the agents decide that they need to better adapt to new situations. To build them, bigger, larger, and more diverse volumes of data are needed.

That doesn't mean that, in the entire process of creating these intelligent agents, there is no human element (not-so-autonomous). Rather, if you want a system to autonomously recognize dogs you need humans to tell the system: this is a dog. You also need a lot of expertise to put the right machine learning method in place for a given domain of application, to test the resulting models, and to fine-tune it before deploying the final intelligent agent in its operating environment.

3
AI IN DESIGN

The fact that more and more systems have integrated some type of intelligent agent in their construction has opened up a wide range of intersections between the goals, tasks, and methods of design and artificial intelligence. The following highlights some of the areas in which AI is applied in design.

Materials can be understood as something passive and a given or, conversely, something that can be designed to exhibit a set of desired properties; potentially, materials can be seen as something that can make these properties evolve over time in response to an environment. Exploring the vast combinatorial space of the ways to combine existing matter or modify existing materials is one area where the use of artificial intelligence seems quite appropriate. Searching for different combinations and operations to modify individual elements or their combinations corresponds to a very old configuration of AI, that is, the one that understands intelligent behavior as the smart exploration of a search space using heuristics (Russell and Norvig 2003). One can trace some of this view on the application of AI to the creation of new materials back to the 80s.

More recently, one can also find the use of the "second wave" of artificial intelligence systems in the creation of new materials. That is, the typical cycle of data gathering, machine learning application, testing a model, and applying it to the task at hand. In this case, repositories of material properties and processes are used to come up with a possible set of combinations and procedures to obtain the desired performance of a material substrate that can be a composite or a more fundamental change, for example, pressure, thermal response, or elasticity. At a different scale, other properties of materials can be explored and generated through AI methods such as the texture of a surface and also to embed programmability in the materials themselves (Ion et al. 2017).

AI is being used in product design by extending its procedures at the level of the product. For example, the work of Troy Nachtigall explores the connection between data and materialities (Nachtigall et al. 2019). By using data about materials, properties, user behavior, and other sources, a system that learns from the interconnection of all of them is used to derive an ultra-personalized product. The models followed in design could also be the input of a learning system (Tucker 2016).

UX and HCI are likely the subdisciplines in which the most discussion and use of AI and ML methods has taken place. There are many descriptions of systems that have used some type or another of AI and ML to improve several aspects of interaction and HCI. The areas of personal assistants and interface design are very active in that sense. The classical conceptualization hinged on assumptions about the role of users' cognitive tasks, symbolic models of planning which, interestingly enough, had similarities to "Classic" or "First Wave" AI (up to the 90s) (Suchman 2007). However, the behavior of machine-learning systems is done through complex numerical models of the world rather than symbolic descriptions, mediating the interaction of users with the world (Blackwell 2015).

The introduction of AI and ML in these areas has also sparked a very interesting discussion about which methods for design research should be used. The answer to this stems from the adaptation of current trends, such as user-centered design (Google Design), to other more innovative proposals that recognize the agency of the new materials and their different roles with respect to human users such as integration vs. interaction (Farooq 2016). The ability of intelligent agents to predict in order to adapt to the human (or non-human) user/agent has made some propose that such interaction should be understood as a variation of a general "anticipatory design" (Van Bodegraven 2017).

Putting human users and artificial agents on par, or nearly so, in terms of agency results in understanding and designing systems, services, and products "from the point of view of the agent" (Cruickshank 2017). This calls for new design methods. For example, "Thing ethnography" (Giaccardi, Cila, Speed, and Caldwell 2016a, 2016b) and animistic design research (Marenko and van Allen 2016) could be seen as approximations adopting a non-anthropocentric point of view of design (DiSalvo 2011). They try to come up with actionable methods within a post-human approach to design.

4 CHALLENGES

The overlapping of AI, Machine Learning, Big Data, and design has opened up new possibilities for design. Also, it has revealed new difficulties in different aspects of design. Without being in any way exhaustive, we discuss several challenges below.

Assuming artificial

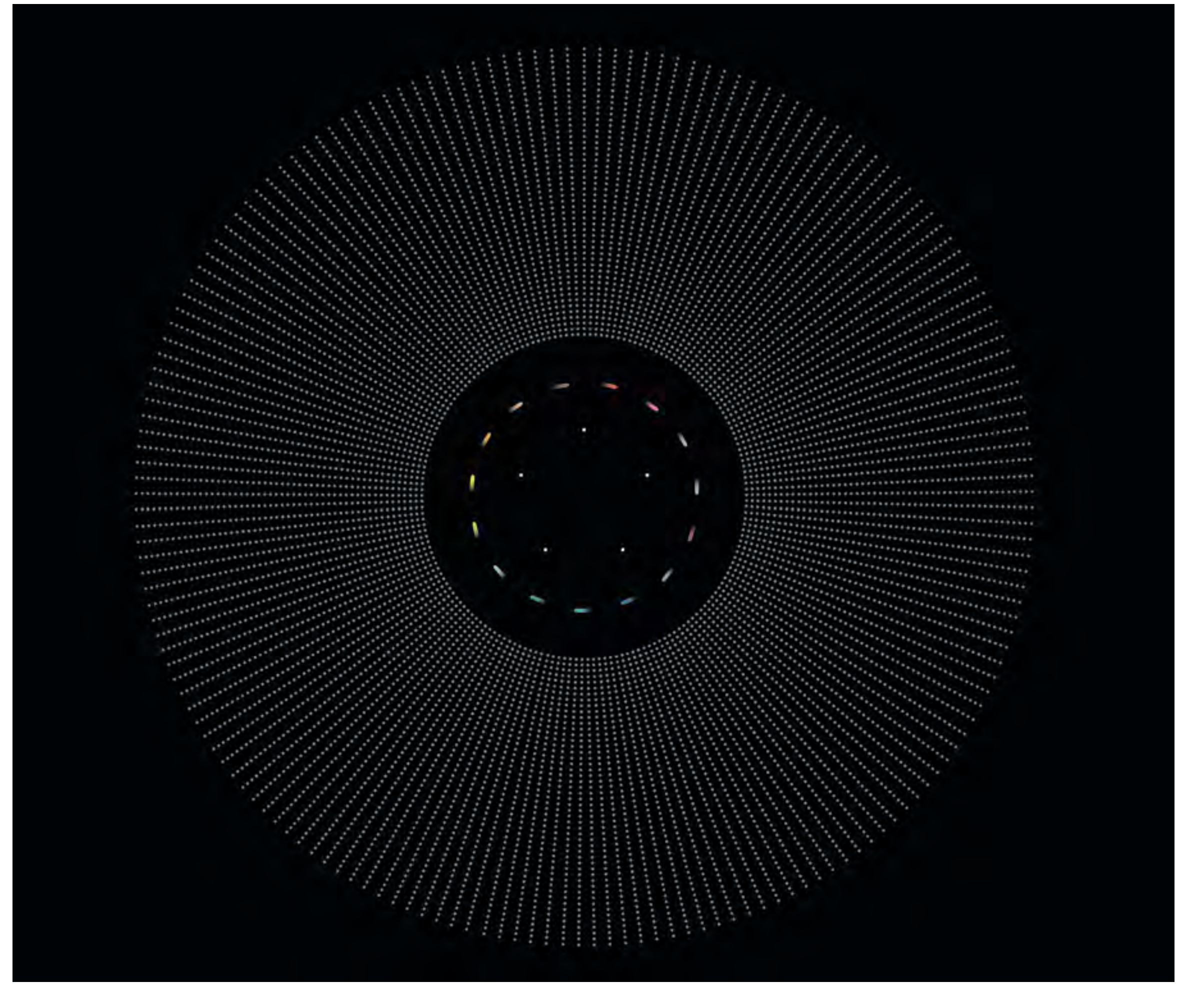
Going beyond the assumption that all users are necessarily human and that the level of agency of designed things is much greater than before puts into question some of the most frequent approaches in design. User-centered design is probably the methodology that is farthest from this assumption. Nevertheless, it is used by teams creating AI systems (Google Design). Some researchers in design start from that point and recognize the dynamic, evolutive nature of systems designed with AI (or for AI). In particular, learning creates new relationships of co-adaptation between users and systems (Leahu 2016). Recognizing the ongoing and changing relationship between the capabilities of human and artificial agents is the basis for approaches such as co-performance (Kuije and Giaccardi 2018), which questions the usual distribution of agency between human and artificial agents that are part of a system. Actually, both of them learn and change because they co-perform and co-evolve. As mentioned before, other approaches such as animistic design could also share this understanding of the mutual evolution through learning in mixed intelligent populations of agents (human or artificial) (van Allen 2013).

Connecting AI, ML, data and design cultures

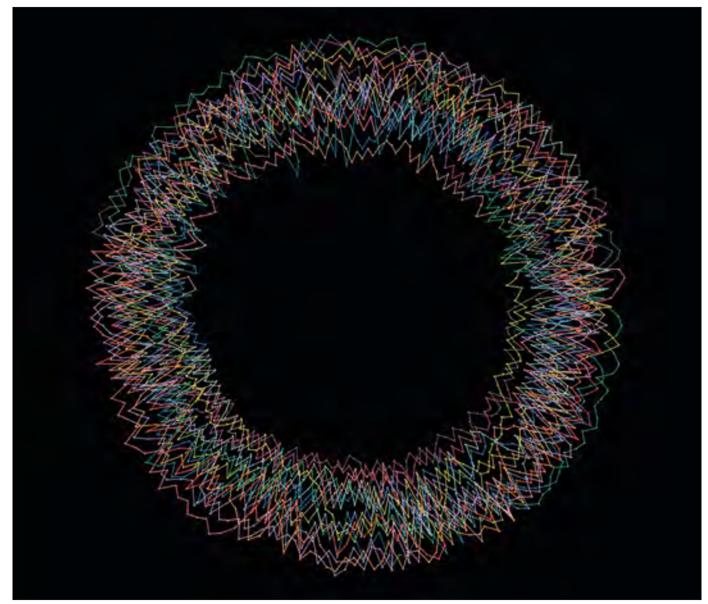
As described above, the present relationship between data, machine learning, and artificial intelligence agents is closer than ever. This brings new requirements for design professionals. There is a need to get the right data for the intended goals of the system being designed. This data should somehow ensure that the resulting model is of quality and actionable. However, the current methods of Machine Learning in particular and Artificial Intelligence in general evaluate models from the perspective of data science and are not directly translatable to other evaluation criteria used in design and that have to do with more qualitative aspects, for example, the quality of the user experience, to name one (Dove 2017; Kuniavski 2017). There is a need to get the right data, the right ML method, and the right resulting model for the task at hand and it is still difficult to anticipate how this could impact design attributes. Studies of actual use of data for AI and Design show a rift between what is the usual practice in data engineering and machine learning with respect to design (Yang 2017). For the time being, it seems that the only way to proceed is through the creation of multidisciplinary teams (Girardin and Lathia 2017).

Operationalizing new understandings of established concepts in design

In dealing with systems composed by one or several artificial agencies, new situations arise for human users that demand an expansion of current design concepts. Usability is stressed under new demands. For example, understanding the behavior of an artificial intelligent agent with which a user interacts, and specifically an agent that learns, stresses the current approach based on plans and cognition (Suchman 2007; Van Allen 2017). Understanding why the agent made a decision to behave in a given way involves new aspects (Seshadri 2017; Huang 2017). Answers to these questions coming from the technical side (whether from ML, AI or Data Systems) are complex and difficult to communicate to the user in an effective, simple, and unobtrusive way. Current methods are divided into methods that resort to transparency (Kizilcec 2016; DTL 2018; Sangüesa 2018) or explanation (Gunning 2018). This lack of understanding creates lower quality user experience (Springer 2017; Kuniavsky 2017) and a lack of user confidence with respect to the intelligent system. Nevertheless, this is especially difficult in systems derived by an application of neural models, in particular Deep Learning (Lei 2016). This puts trust under attack given the known problematic implications of AI, Machine Learning, and Big Data: the transmission and amplification of bias (Eubanks 2018; Crawford 2016) and the reinforcement of discrimination (Eubanks 2018; Sweeney 2013; FATML 2017, 2018). The creation of trustworthy AI systems is still an open debate but definitely connects with the values and practices of design (Floridi 2019; IEEE 2018). The ethical design of intelligent systems requires combining these aspects with responsibility (Baylé 2019) and ethical expertise (Casacuberta and Guersenzvaig 2018). Last but not least, ethics for intelligent agents – technological systems based on an infrastructure that is environmentally quite demanding – should be understood in an open way, involving other entities and not just humans in their con-



Plataforma Experiencial by Cristina Melcior. Elisava Graphic Design Final Degree Project, 2016.



sideration, that is, there is also a connection with efforts to design in a sustainable way in this area (DiSalvo 2010). These sustainability efforts must be extended to reduc the environmental impact that the development of AI systems generates. Recent research shows that the carbon footprint generated by training common large AI models is almost five times larger than the average American car, including the car itself and the fuel it consumes (Hao 2019). The numerical series present in patterns may perhaps belong to a Platonic realm unaffected by sustainability concerns. The hardware that makes machine learning possible, however, has very real environmental effects. Developing more sustainable hardware is an ethical imperative.

All design has a normative dimension. From speedbumps that make us slow down when we are driving a car to forks or spoons that roughly determine the amount of food that we put in our mouths. Designed artefacts contain and transmit norms and standards. As philosopher of technology Carl Mitcham (1995) writes, "different designs embody (implicitly or explicitly) distinct sociopolitical assumptions and visions of life, designing itself constitutes a new way of leading, or a leading into, different technological lifeworlds." Even if the actual use of artefacts is never fully determined in advance by the designer, their decisions can and do have important ethical consequences. Simply by being concerned by "how things ought to be," as Herbert Simon (1961) famously expressed, designers undertake an ethically challenging task for which they need to be prepared and aware of. This becomes the case especially when they are not the only agencies involved in actions that are vested with ethical relevance.

5 THIS ISSUE

Temes de Disseny felt that there was room to invite reflection regarding this subject. We opened up a call to researchers and practitioners in design to share their explorations and realizations that connected Artificial Intelligence, Machine Learning, Big Data, Intelligent materials, and other emerging technologies. We are happy to share with you the papers resulting from a thorough peer review selection process.

Landscape Design Methodology: pattern formation through the use of cellular automata by Sergi Abellán, Marcel Bilurbina, and Marilena Christodoulou shows how to apply a well-established artificial construct in computing, cellular automata, to help in the design of landscapes. Cellular automata can be seen as a collection of agents with simple rules that act as their "knowledge." Nevertheless, collections of these automata typically exhibit very complex emergent behaviors that replicate the behaviors of colonies of living organisms. The comparison of the evolution of the artificial and natural counterparts shows a generative intelligent approach in the domain of landscape design but also points to a more general applicability.

Designing Predictive Tools for Personalized Functionalities in Knitted Performance Wear by Martijn ten Bhömer, Hai-Ning Liang, Difeng Yu, Yuanjin Liu, Yifan Zhang, Eva de Laat and Carola Leegwater, is an exploration of the possibilities of Industry 4.0 and how it impacts the creative design process. The combination of data, machine learning, and simulation brings new possibilities for extraordinary levels of personalization in textile manufacturing and in knitwear in particular. This article is a case study in this domain that explores how the predictive software inspired by machine learning can help in the processes. They focus in greater depth on the characteristics of this software that can help in the creativity of the designers involved, with a special view on the design of interactions and interfaces.

Crafting Soft Wearables, with and through digital technologies by Bruna Goveia, Kristina Andersen, and Oscar Tomico reviews and reflects on the last seven years of the Wearable Senses lab's operation. The paper focuses on a view of intelligent wearables that takes them as a fully-programmable entity. The authors extend their reflection to the entire ecosystem around wearables, taking into account the different levels of personalization, the manufacturing processes, and the relationship with related services, users, and other stakeholders. In doing so, special interest is placed on the role of data and computing in all the processes.

Practicing Fashion with the Anthropocene by Patricia Wu approaches the practice of design from a speculative perspective where there is a displacement of the human-centric focus, a phenomenon that characterizes the Anthropocene. The author explores the consequences of this shift in agency in which the world is populated by non-human, highly autonomous entities, or "Odradeks." Through the description of the author's practice-based process, the possibility of reaching a fashion design practice that fully embodies an ecological, non-human-centric view is described.

The biological encoding of design and the premises for a new generation of "living" products: the example of Sinapsi by Sabriana Lucibello and Carmen Rotondi describes the process by which they created an intelligent device to help blind people navigate through different environments. In order to design in a world full of artificial intelligences, they used biological models to steer their design process. The authors argue that this helps improve the quality of life of the user of an intelligent system and promotes human qualities as well. The fact that they tested their approach in a setting where the co-dependence of human users and artificial autonomous systems is so high reinforces the interest in their findings as they could point a way towards addressing other intelligent systems in other domains.

*Designing organs at the Transpecies Society: hybrid practices between cybernetic*s *and artificial intelligence* by Tatiana Afanador and Judit Parés. The Transpecies Society is an association that embraces the cybernetic and the artificial as their main material and the primary focus of their practice. The authors discuss, in a very practical manner, their work methods and how they are related to a culture based on data in a critical relationship with the tenets of Cybernetics and intelligent systems. They discuss their own practices by examining, for example, what goes on in the design and operation, or life, of cyborgs.

The Service Design Platform for People with Dementia: Person-centred Reminiscence Therapy with Artificial Intelligence in Immersive Environments by Jinyoung Lee explores the interconnection of design and virtual environments to facilitate reminiscence therapy for people with dementia and their caregivers. The system is centered around the storage, categorizaton, and retrieval of personal memories for patients. The changes introduced by the use of Virtual Reality environments in this regard are considered since, the author argues, they could expand the current way memories are represented and relived. The ethical aspects of these changes are also considered.

Death Inc. by the Domestic Data Streamers design collective. This pictorial shows an exhibit that was presented at the "Design Does" exhibition. This exhibition took place at The Design Museum of Barcelona and revolved around the different dimensions of contemporary design and its importance in society. Death Inc. was an installation about killer robots to spark debate about how Autonomous Weapon Systems failed the Principle of Discrimination. It referred to the Samsung SGR-A1 model, currently used in the demilitarized zone of the South Korea-North Korea border. This model, considered the first of its kind, is able to select and shoot human targets without requiring a person to authorize the operation.

A framework for systematically applying humanistic ethics when using AI as a design material by Richelle Dumond, Kyle Dent, and Mike Kuniavsky deals with the hard task of coming up with a consistent and comprehensive set of guidelines and practices to help designers confront the ethical dilemmas of their practice when dealing with the creation of intelligent systems. They place special focus on the difficulties brought about by the new levels of autonomy in these types of systems and how to tackle them to uphold human rights, respect individuals' privacy, ensure personal data protection, and enable freedom of expression and equality, a set of conditions for maintaining a humanistic ethics.

We would like to thank all authors for their contributions and the reviewers for their efforts. We hope that this issue helps designers to position themselves with respect to AI and emerging technologies and brings about the best possible outcomes from their research and practice.

BIOGRAPHIES

Ramon Sangüesa
Technical University of Catalonia

Ramon Sangüesa is a professor at the Technical University of Catalonia (UPC). Previously, he was the director of the technology research area of Elisava's design and engineering school as well as a coordinator of the Data Transparency Lab, an international consortium that researches data transparency techniques. He is the founder of MyData Hub Barcelona within the mydata.org network that promotes sovereignty over personal data. He is an affiliate researcher at the Center for Organizational Innovation with the Department of Sociology at the University of Columbia in New York.

Ariel Guersenzvaig
Elisava School of Design and Engineering

Dr. Guersenzvaig is the director of the Master's Degree in Design and Communication, an official Master's Degree that gives access to doctoral studies. He also directs a professional program, the Master's in Design and Project Management for the Internet. He is a graduate and postgraduate professor at Elisava and a visiting professor at various universities (UPF, UPC). He has a PhD in design theory from the University of Southampton (UK), an MA in ethics from the University of Birmingham (UK), a Diploma of Advanced Studies in design research from the University of Barcelona (ES), and a Degree in Publishing from the University of Applied Sciences of Amsterdam (NL).

REFERENCES

Agarwal, P. 2019. "Redefining Banking and Financial Industry through the application of Computational Intelligence". In *2019 IEEE Advances in Science and Engineering Technology International Conferences (ASET)*, 1-5. DOI: 10.1109/ICASET.2019.8714305.
AI Expert Newsletter: *W is for Winter*. Archived 9 November 2013 at the Wayback Machine.
Alpaydin, E. 2016. *Introduction to machine learning*. Cambridge, Mass.: MIT Press.
Baylé M. Designing Responsibly AI. 2019. Master's Thesis: Experience Design MA, Hyper Island (Manchester). Downloadable from https://www.designingresponsiblywithai.com/
Blackwell, A.F. 2015. "Interacting with an inferred world: the challenge of machine learning for humane computer interaction". *In Proceedings of The Fifth Decennial Aarhus Conference on Critical Alternatives* (CA '15). Aarhus University Press, 169-180. DOI: 10.7146/aahcc.v1i1.21197
Bond, R. M., C.J. Fariss, J.J. Jones, A.D.I. Kramer, C. Marlow, J.E. Settle, and J.H. Fowle. 2012. "A 61-Million-Person Experiment in Social Influence and Political Mobilization". Nature 489(7415):295–298.
Booth, R. 2014. "Facebook reveals news feed experiment to control emotions". *The Guardian*, June 30, 2014. https://www.theguardian.com/technology/2014/jun/29/facebook-users-emotions-news-feeds
Brooks, R. A. 1991. "Intelligence Without Reason". *In Proceedings of 12th International Joint Conference on Artificial Intelligence*, 569-595. San Francisco: Morgan Kaufman.
Casacuberta, D. and A. Guersenzvaig. 2019. "Using Dreyfus' legacy to understand justice in algorithm-based processes". *AI & Society* 34(2):313-319. DOI: 10.1007/s00146-018-0803-2.
Cramer, H., and J. Thom. 2017. "Not-So-Autonomous, Very Human Decisions in Machine Learning: Questions when Designing for ML". *AAAI Spring Symposium Series*. Palo Alto, CA : AAAI Publications.
Crawford, K. 2016. "Artificial Intelligence's White Guy Problem". *The New York Times*, June 25, 2016. https://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html
Cruickshank, L., and N. Trivedi. 2017. "When Your Toaster is a Client, how do you design? Going Beyond Human Centred Design". *The Design Journal*, 20(sup1):S4158-S4170. DOI: 10.1080/14606925.2017.1352914
Davis, Z.S. 2019. *Artificial Intelligence on the Battlefield : An Initial Survey of Potential Implications for Deterrence, Stability, and Strategic Surprise*. Center for Global Security Research Lawrence Livermore National Laboratory. https://cgsr.llnl.gov/content/assets/docs/CGSR-AI_BattlefieldWEB.pdf
Dignum, V. 2018. "Ethics in artificial intelligence: introduction to the special issue". *Ethics and Information Technology* 20(1):1–3.
DiSalvo, C., and J. Lukens. 2011. "Nonathropocentrism and the Nonhuman in Design: Possibilities for Designing New Forms of Engagement with and through Technology". In *From Social Butterfly to Engaged Citizen: Urban Informatics, Social Media, Ubiquitous Computing, and Mobile Technology to Support Citizen Engagement*. Marcus Foth, et al. (eds.), 421-436. Cambridge, MA: MIT Press.
DiSalvo, C., P. Sengers, and H. Brynjarsdóttir. 2010. "Mapping the landscape of sustainable HCI". In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '10), 1975-1984. New York: ACM. DOI: 10.1145/1753326.1753625.
Dourish, P. 2017. The Stuff of Bits: *An Essay on the Materialities of Information*. Cambridge, Mass.: MIT Press.
Dove, G., K. Halskov, J. Forlizzi, and J. Zimmerman. 2017. "UX Design Innovation: Challenges for Working with Machine Learning as a Design Material". In *Proceedings of the 2017* CHI Conference on Human Factors in Computing Systems, 278-288.
DTL. 2018. *Data Transparency Lab*. http://datatransparencylab.org
Eubanks, V. 2018. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York: St. Martin's Press.
Farooq, U, and J. Grudin. 2016. "Human-computer integration". *Interactions* 23(6):26-32. DOI: 10.1145/3001896
Floridi, L. 2019. "Establishing the rules for building trustworthy AI". *Nature Machine Intelligence*. DOI: 10.1038/s42256-019-0055-y.
Giaccardi, E., N. Cila, C. Speed, and M. Caldwell. 2016a. "Thing Ethnography: Doing Design Research with Non-Humans". In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems (DIS '16)*, 377-387. New York: ACM. DOI: 10.1145/2901790.2901905.
Giaccardi, E., C. Speed, N. Cila, and M. Caldwell. 2016b. "Things as Co-ethnographers: Implications of a Thing Perspective for Design and Anthropology". In *Design Anthropological Futures. R. C. Smith, et al. (eds.)*, 235-248. London: Bloomsbury
Girardin, F., and N. Lathia. 2017. "When User Experience Designers Partner with Data Scientists". *AAAI Spring Symposium Series*. Palo Alto, CA: AAAI Publications.
Goodfellow, I., Y Bengio, and A. Courville. 2016. *Deep Learning*. Cambridge, MA: MIT Press.
Goodfellow, I., J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. 2014. "Generative Adversarial Networks". In *Proceedings of the International Conference on Neural Information Processing Systems (NIPS 2014)*, 2672–2680. Cambridge, MA: MIT Press.
Grudin, J. 2009. "AI and HCI: Two Fields Divided by a Common Focus". *AI Magazine* 30(4):48-57. DOI: 10.1609/aimag.v30i4.2271.
Gunning, David. 2018. *Explainable Artificial Intelligence*. DARPA. Accessed February 4th, 2019: https://www.darpa.mil/program/explainable-artificial-intelligence
Hao, K. 2019. "Training a single AI model can emit as much carbon as five cars in their lifetimes". *MIT Technology Review*. https://www.technologyreview.com/s/613630/training-a-single-ai-model-can-emit-as-much-carbon-as-five-cars-in-their-lifetimes.
Hogarth, I. 2018. "AI Nationalism". *Ian Hogarth*. https://www.ianhogarth.com/blog/2018/6/13/ai-nationalism.
Huang, C., K.H. Ting-Hao, and K. Lun-Wei. 2017. "Challenges in Providing Automatic Affective Feedback in Instant Messaging Applications". In The AAAI 2017 Spring Symposium on Designing the *User Experience of Machine Learning Systems.* https://aaai.org/ocs/index.php/SSS/SSS17/paper/view/15349.
IEEE. 2018. *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems (EADv2)*. IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead_v2.pdf
Ion, A., L. Wall, R. Kovacs, and P. Baudisch. 2017. "Digital Mechanical Metamaterials". In Proceedings of the 2017 CHI *Conference on Human Factors in Computing Systems (CHI '17)*, 977-988. New York: ACM. DOI: 10.1145/3025453.3025624
Kizilcec, R.F. 2016. "How Much Information?: Effects of Transparency on Trust in an Algorithmic Interface". In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2390– 2395. New York: ACM Press. DOI: 10.1145/2858036.2858402
Koch, J. 2017. "Design implications for Designing with a Collaborative AI". *The AAAI 2017 Spring Symposium on Designing the User Experience of Machine Learning Systems*. https://aaai.org/ocs/index.php/SSS/SSS17/paper/view/15382.
Koch, F, A. Koster, D. Riaño, S. Montagna, N. Shucmacher, A. ten Teije, C. Guttman, M. Reichert, I. Bichindaritz, P. Herrero, R. Lenz, B. López, C. Marling, C. Martin, S. Montani, and N. Wiratunga, eds. 2019. Artificial Intelligence in Health : First International Workshop, AIH 2018. Cham, Switzerland : Springer.
Krishnan, A. 2009. Killer Robots: Legality and Ethicality of Autonomous Weapons. London: Routledge.
Kuije, L. and E. Giaccardi. 2018. "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 (CHI '18). New York: ACM. DOI: 10.1145/3173574.3173699.
Kuniavsky, M., E. Churchill, and M.W. Steenson, eds. 2017. *Papers form the AAAI 2017 Spring Symposium on Designing the User Experience of Machine Learning Systems*. https://www.aaai.org/Library/Symposia/Spring/ss17-04.php
Kückelhaus, Markus, and G. Chung. 2018. A*rtificial Intelligence in Logistics: a collaborative report by DHL and IBM on implications and use cases for the logistics industry*. Available at https://www.logistics.dhl/content/dam/dhl/global/core/documents/pdf/glo-ai-in-logistics-white-paper.pdf
Latour, B. 2005. *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford: Oxford University Press.
Leahu, L. 2016. "Ontological Surprises: A Relational Perspective on Machine Learning". In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems (DIS '16)*, 182-186. New York: ACM. DOI: 10.1145/2901790.2901840
Lei, T., R. Barzilay, and T. Jaakkola. 2016. "Rationalizing neural predictions". arXiv preprint, available at https://arxiv.org/abs/1606.04155.
Lighthill, J. 1973. "Artificial Intelligence: A General Survey". In *Artificial Intelligence: a paper symposium*. [London]: Science Research Council.
Mitcham, C. 1995. "Ethics into design". In *Discovering Design: Explorations in Design Studies*. R. Buchanan, and V. Margolin (eds.), 173-189. Chicago: University of Chicago Press.
Marenko, B., and P. van Allen. 2016. "Animistic Design: How to Reimagine Digital Interaction between the Human and the Non-human." *Digital Creativity 27*(1): 52–70. DOI: 10.1080/14626268.2016.1145127.
Marinchak, C. M., E. Forrest, and B. Hoanca. 2018. "Artificial Intelligence: Redefining Marketing Management and the Customer Experience". *International Journal of E-Entrepreneurship and Innovation (IJEEI)*, 8(2):14-24. DOI: 10.4018/IJEEI.2018070102.
McCarthy, J., M. Minsky, N. Rochester, and C. Shannon. 1955. "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence". http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html
Nachtigall, T., O. Tomico, R. Wakkary, and P. van Dongen. 2019. "Encoding Materials and Data for Iterative Personalization". In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. New York: ACM. DOI: 10.1145/3290605.3300749.
NSTC. 2016. *The National Artificial Intelligence Research And Development Strategic Plan*. National Science and Technology Council; Networking and Information Technology Research and Development Subcommittee. https://www.nitrd.gov/PUBS/national_ai_rd_strategic_plan.pdf
Pedersen, T., C. Johansen, and A. Jøsang. 2018. "Behavioural Computer Science: an agenda for

combining modelling of human and system behaviours". *Human-centric Computing and Information Sciences* 8:7. DOI: 10.1186/s13673-018-0130-0.
Portugal, I., P. Alencar, and D. Cowan. 2017. "The use of machine learning algorithms in recommender systems: A systematic review". Expert Systems with Applications 97:205-227. DOI: 10.1016/j.eswa.2017.12.020.
Rozendaal, M. 2016. "Objects with intent: A new paradigm for interaction design". *Interactions 23*:62-65. DOI: 10.1145/2911330
Russell, S.J., and P. Norvig. 2003. *Artificial Intelligence: A Modern Approach, 2nd ed*. Upper Saddle River, NJ: Prentice Hall.
Sangüesa, Ramon. 2018. "Inteligencia artificial y transparencia algorítmica : 'It's complicated'". *BiD: textos universitaris de biblioteconomia i documentació 41*. DOI: 10.1344/BiD2018.41.12.
Serra, A. 1992. "Design Culture : estudi etnogràfic dels projectes de recerca de la School of Computer Science de Carnegie Mellon University,un campus nord-americà altament informatitzat." PhD diss., Universitat de Barcelona.
Siddike, Md. A., J. Spohrer, H. Demirkan, and Y. Kohda. 2018. "A Framework of Enhanced Performance: People's Interactions With Cognitive Assistants". *International Journal of Systems and Service-Oriented Engineering 8*(3):1-17. DOI: 10.4018/IJSSOE.2018070101.
Simon, H. 1961. *The Sciences of The Artificial*. Cambridge MA: MIT Press.
Springer, A., V. Hollis, and S. Whittaker. 2017. "Dice in the Black Box: User Experiences with an Inscrutable Algorithm". In *AAAI 2017 Spring Symposium on Designing the User Experience of Machine Learning Systems*. Palo Alto, CA: AAAI Publications.
Suchman, L. 2007. *Human-Machine Reconfigurations: Plans and situated actions.* New York: Cambridge University Press.
Turing, A. 1950. "Computing Machinery and Intelligence". Mind LIX(236):433–460. DOI: 10.1093/mind/LIX.236.433.
Turkle, S. 2006. "Artificial Intelligence at Fifty: From Building Intelligence to Nurturing Sociabilities." Paper presented at the *Dartmouth Artificial Intelligence Conference*, Hanover, NH, USA, July 15, 2006. http://www.mit.edu/~sturkle/ai@50.html.
van Allen, P., J. McVeigh-Schultz, B. Brown, H.M. Kim, and D. Lara. 2013. "AniThings: animism and heterogeneous multiplicity". In CHI'13 *Extended Abstracts on Human Factors in Computing Systems, 2247-2256*. Paris: ACM.
van Allen, P. 2017. "Reimagining the Goals and Methods of UX for ML/AI". In AAAI 2017 *Spring Symposium on Designing the User Experience of Machine Learning Systems*, 431-434. https://aaai.org/ocs/index.php/SSS/SSS17/paper/view/15338.
van Bodegraven, J. 2017. "How Anticipatory Design Will Challenge Our Relationship with Technology". In *AAAI 2017 Spring Symposium on Designing the User Experience of Machine Learning Systems, 435-438*. https://aaai.org/ocs/index.php/SSS/SSS17/paper/view/15352.
Verbeek, P. 2011. *Moralizing Technology: Understanding and Designing the Morality of Things*, Chicago: University of Chicago Press.
Villani, C. 2018. For a *Meaningful Artificial Intelligence : Towards a French and European strategy*. Accessed March 19, 2018. https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf
Webster, G., R. Creemers, P. Triolo, and E. Kania. 2017. "China's Plan to 'Lead' in AI: Purpose, Prospects, and Problems". *New America*, August 1, 2017. https://www.newamerica.org/cybersecurity-initiative/blog/chinas-plan-lead-ai-purpose-prospects-and-problems/
Wilson, H.J., P.R. Daugherty, and N. Morini. 2017. "When AI Becomes the New Face of Your Brand". *Harvard Business Review*. June 27, 2017. https://hbr.org/2017/06/when-ai-becomes-the-new-face-of-your-brand
Winograd, T. 2006. "Shifting viewpoints: Artificial intelligence and human–computer interaction". *Artificial Intelligence 170*(18):1256-1258. DOI: 10.1016/j.artint.2006.10.011.
Wooldridge, M. 2002. *An Introduction to MultiAgent Systems*. Chichester: Wiley.
Yaneva, A. 2009. "Making the Social Hold: Towards an Actor-Network Theory of Design." *Design and Culture 1*(3):273-288. DOI: 10.1080/17547075.2009.11643291.
Yang, Q. 2017. "The role of design in creating machine-learning-enhanced user experience." In The *AAAI 2017 Spring Symposium on Designing the User Experience of Machine Learning Systems, 406–411*. https://aaai.org/ocs/index.php/SSS/SSS17/paper/view/15363
Zuboff, S. 2015. "Big other: surveillance capitalism and the prospects of an information civilization". *Journal of Information Technology 30*(1):75-89.

R. Sangüesa i A. Guersenzvaig

La intel·ligència artificial com a material de disseny: treballar amb noves agències

Traducció al Català

PARAULES CLAU

Intel·ligència Artificial, Disseny, Materials, Agència, Aprenentatge Automàtic, Ètica.

1
INTRODUCCIÓ

Intel·ligència artificial (IA) s'ha convertit en una expressió d'ús domèstic, una expressió en voga en els mitjans de comunicació, però la IA també és una indústria en si mateixa i un conjunt de pràctiques associades a un negoci i a uns objectius professionals. Com a tecnologia simbòlica general que ha rebut enormes inversions, ha influït en molts àmbits d'activitat. Els *chatbots* poblen Internet, on s'utilitzen com a objectius polítics (Bond et al. 2012) i de màrqueting (Wilson 2017). Per la seva banda, els sistemes intel·ligents organitzen les cadenes logístiques (Kückelhaus i Chung 2018), dirigeixen l'atenció de la gent cap a determinats productes (Marinchak et al. 2018), influeixen en la seva pròxima compra (Portugal et al. 2017), modifiquen les seves emocions (Booth 2014), interpreten proves mèdiques (Koch et al. 2018), interactuen de noves maneres amb usuaris (Siddike et al. 2018), decideixen qui rep una beca o un préstec (Agarwal 2019; Eubanks 2018) o quan i com lliurar una guerra (Davis 2019). D'una manera o d'una altra, els sistemes intel·ligents classifiquen, programen, planifiquen i creen productes, serveis i experiències. Prenen decisions contínuament en molts aspectes de la nostra vida, i així contribueixen a donar-li forma (Turkle 2006).

No és la primera vegada, però, que es produeix un auge en l'aplicació de tècniques d'IA en la indústria. Durant els anys vuitanta i noranta es va produir una explosió d'aplicacions d'aquesta tecnologia que van despertar grans esperances, però la fragilitat que va mostrar l'ús real de les tecnologies de l'època va acabar instaurant el primer "Hivern de la IA" (AI Expert Newsletter; Russell i Norvig 2003). Actualment, el nivell de digitalització en amplis sectors d'activitat i la millora en moltes tècniques d'IA han creat el terreny propici per a un nou ús generalitzat de la IA. Mentre que les versions antigues d'IA necessitaven l'ús dels denominats sistemes experts, amb un alt cost de desenvolupament i un alt nivell de dificultat d'entrenament, l'actual combinació de volums i disponibilitat de dades de molts tipus d'aplicacions ha facilitat l'ús de tècniques d'aprenentatge automàtic, que històricament era una subdisciplina de la IA. L'aprenentatge automàtic ha accelerat la construcció i l'entrenament de sistemes d'IA i ha superat l'antiga escassetat de recursos a partir dels quals es podia extraure coneixement i amb els quals els sistemes d'IA podien dur a terme les seves tasques.

Actualment existeix la idea que la IA tindrà un paper determinant a les nostres societats. Es considera un actiu estratègic que, per exemple, afectarà la competència internacional (Villani 2018; Webster et al. 2017; NSTC 2016; Hogarth 2018). S'ha atribuït a la IA una gran influència i, per bé o per mal, se li ha confiat l'organització d'aspectes que van des del comportament personal fins a les estratègies econòmiques. Les seves connexions amb l'economia conductual (Pedersen 2018), la vigilància digital (Zuboff 2015) i les aplicacions militars (Krishnan 2009) fan que la preeminència de la IA com a orquestradora de la vida sigui un mal presagi i han donat lloc a una intensa activitat entorn de les seves implicacions ètiques (Floridi 2019; Dignum 2018; Casacuberta i Guersenzvaig 2018).

En la seva accepció de pràctica professional i sector en si mateix, el disseny també ha experimentat les friccions i l'entusiasme que desperta la IA. Els professionals de diverses àrees del disseny s'han vist involucrats en la IA de diverses maneres. Potser els primers que van percebre l'impacte de la IA van ser els professionals del disseny que treballaven en l'àmbit de la interacció persona-ordinador (HCI, per les seves sigles en anglès) i de l'experiència d'usuari (UX) (Winograd 2006; Grundin 2009), que històricament ha estat el terreny en què allò digital ha tingut més impacte en el disseny. Les tècniques d'IA, però, i especialment les tècniques d'aprenentatge automàtic, també han fet incursions en altres àmbits del disseny.

Molts plantejaments generatius del disseny utilitzen profusament la interconnexió entre les dades i els algoritmes, especialment els d'aprenentatge automàtic, per idear i produir nous productes i ho fan de manera que l'algoritme intel·ligent assumeix un paper que és alguna cosa més que el paper passiu o reactiu d'una eina a mans dels dissenyadors (Koch 2017). A més a més, treballar amb serveis de recomanació i assistents intel·ligents ha comportat nous problemes i objectius per al disseny. A poc a poc, ens vam adonar que, en realitat, alguns dissenys no es creen amb la intel·ligència artificial o per la intel·ligència artificial, sinó per a aquest tipus d'entitats (DiSalvo 2011). Els sistemes d'IA, doncs, de vegades fan alhora la funció d'eina, de col·laborador i d'usuari.

Ens va semblar que aquesta situació mereixia una reflexió des del punt de vista del disseny i de l'enginyeria del disseny, i vam buscar un tema comú que fos útil per analitzar la relació entre la IA i el disseny, així com altres tecnologies emergents com el Big Data o els materials intel·ligents. La IA es considera bàsicament una entitat "soft" però, tal com passa amb altres àrees de la ciència de la informació i la computació, també té un component material (Dourish 2017, 33-59). Vam pensar que aquest podia ser un bon punt de partida per contrastar diferents interaccions entre la IA i el disseny, i alhora, per començar a aclarir l'estat actual de les esmentades connexions. En fer-ho, però, també volíem destacar que com a material de disseny, la IA té característiques molt peculiars. És difícil interpretar-la com un material passiu i constant. Per contra, no només és mal·leable, sinó que té un cert nivell de "vida", o si més no d'autonomia, en si mateixa. Dissenyar amb IA és clarament treballar amb un material que evolucionarà de formes més variades i més difícils de planificar per endavant, si més no en comparació amb altres materials més tradicionals. Treballar amb fusta contraplacada o amb alumini està circumscrit totalment a les lleis de la física, un fet que els converteix en materials altament predictibles. Els algoritmes d'IA no són aliens al terreny de la física, no cal dir-ho, però la seva mateixa naturalesa fa que el seu comportament pugui ser molt complicat de vaticinar detalladament. De moment, no hi ha cap manera directa d'analitzar un model d'IA i fer que sigui totalment predictible.

Existeix algun concepte comú que, seguint el mateix raonament, ens pugui ajudar a estudiar els diferents tipus de disseny que s'estan creant en relació amb la IA, l'aprenentatge automàtic i les dades? Hi ha algun concepte que les dues disciplines utilitzin i que ens pugui ajudar a fixar les relacions que mantenen? Després de reflexionar i debatre durant un temps vam arribar a la conclusió que la noció d'"agència" era una bona candidata a ajudar-nos a dibuixar una connexió pràctica amb els objectes de la IA, l'aprenentatge automàtic, el Big Data i el disseny. Si més no, algunes similituds de conceptualització i ús pràctic que hi havia en totes dues disciplines eren intrigants.

La diferència que la IA pot introduir en general –i en el disseny en particular– es podia interpretar com un salt en el nivell d'agència causat per l'ús de la IA com a material de disseny. En el disseny s'ha introduït i utilitzat l'"agència" en diverses subdisciplines i pràctiques. Un exemple n'és la investigació de disseny, però també s'ha aplicat com a guia en la pràctica del disseny. La noció d'"agència" arrossega certa tradició, per exemple, de camps com la Actor-Network Theory (Latour 2005; Yaneva 2009). En disseny, es reconeix l'agència d'objectes, que pot estar relacionada amb conceptes com les affordances o es pot aplicar en investigació d'usuari. D'alguna manera, l'agència és un concepte acceptat en la pràctica del disseny. Conceptes com els "objectes intencionals" (Rozendaal 2016), per exemple, s'assemblen molt als tipus d'agència que s'utilitzen en IA. Es podria dir que ara la IA, efectivament, és la disciplina que crea agents artificials amb els seus propis objectius, desitjos i intencions. Els acadèmics de la postfenomenologia (per exemple, Verbeek 2011) articulen una perspectiva filosòfica en què l'agència (moral) es redueix a una qüestió de mediació entre persones i tecnologies en lloc de ser una qüestió purament humana.

De fet, en intel·ligència artificial, el concepte d'agència i agències s'ha convertit en un element central i és un constructe fonamental per a la construcció de sistemes complexos d'IA (Wooldridge 2002).

Hi ha diferències entre les conceptualitzacions utilitzades en cada disciplina. Per diferenciar-les podríem retrocedir una mica i analitzar els orígens i els objectius de la IA i per què i com el concepte d'agència té actualment un paper fonamental en diverses disciplines.

2
L'AGÈNCIA EN ELS SISTEMES D'INTEL·LIGÈNCIA ARTIFICIAL

L'inici de la intel·ligència digital com a disciplina se sol associar a la celebració de l'Escola d'Estiu de Dartmouth sobre Intel·ligència Artificial l'any 1956 (McCarthy et al. 1955). Aquella trobada va reunir investigadors en matemàtiques, tecnologia de la informació, lògica, computació i psicologia. Tots compartien un mateix entusiasme per les possibilitats que obrien els ordinadors. Els pioners de Dartmouth van preveure que aquelles màquines podien superar l'ús que havien tingut fins aleshores. Van pensar que els ordinadors podien ser entitats capaces de fer moltes més coses que simplement processar xifres. Van pensar que els ordinadors podrien manipular molts tipus de representacions diferents, incloses representacions del coneixement. El pensament podia ser, tal com ja havia dit Hobbes feia segles, una simple qüestió de càlcul. Tot i que aquesta visió de càlcul de nous tipus de símbols, i no només de xifres, era la dominant en la idea d'intel·ligència de les màquines que va caracteritzar la trobada de Dartmouth, una altra part del mateix grup original estava més interessada en càlculs més simples. Pensaven que no arribarien a la intel·ligència artificial amb estructures simbòliques complexes, sinó replicant la simple combinació de senyals de la infraestructura neuronal dels sistemes intel·ligents cognitius. El grup de Dartmouth va emetre una declaració que, més que un programa detallat d'investigació científica, semblava una llista d'exemples de tasques que associaven al comportament intel·ligent i que pensaven que calia estudiar. Paga la pena fer una ullada a la llista (McCarthy et al. 1955), ja que fa palesa l'ambició del plantejament del grup de Dartmouth i també evidencia els límits de la imaginació d'aquell grup fundacional: tenint en compte la capacitat dels ordinadors en aquella època, calia ser valent per obtenir d'aquelles màquines les tasques extremadament més complexes que necessitarien en el futur per atribuir-los intel·ligència. Era un exercici d'especulació i projecció, potser inspirat per un clima d'investigació amarat dels èxits d'Alan Turing i altres científics en els anys trenta i que van dur Turing a escriure el seu article fundacional "Maquinària computacional i intel·ligència" (Turing 1950). Un dels participants en aquella històrica trobada va definir l'objectiu global de la incipient disciplina de la intel·ligència artificial amb les paraules següents:

"Estudiar com es poden crear sistemes que es comportin d'una manera que, si fossin humans, es pugui considerar 'intel·ligent'". (Minsky 1986)

Curiosament, cadascun dels participants en aquella naixent disciplina tractava l'objectiu comú des de la pràctica de la disciplina de la qual provenia. Més que una metodologia científica –entesa com els mètodes de les ciències dures teòriques o experimentals–, s'estimaven més un plantejament constructiu o, segons una terminologia posterior, un plantejament construccionista. És a dir, sabrien com aconseguir aquells objectius construint sistemes que els assolissin i aprenent del procés. Aquests són bàsicament els trets d'una "cultura del disseny" (Serra 1992), tal com podria acceptar Herbert Simon, un dels més destacats assistents a les trobades de Dartmouth (Simon 1961).

Aquella trobada va donar lloc a moltes activitats i ràpidament es va començar a produir un gran nombre d'èxits inicials. A finals dels anys vuitanta i principis dels noranta, l'àmbit més prometedor de la IA era el de sistemes experts. Aquests es van convertir en sistemes de coneixement i van donar lloc a una disciplina totalment nova: l'enginyeria del coneixement. Els sistemes experts eren bàsicament sistemes de raonament que utilitzaven una representació del coneixement acumulat per experts humans altament qualificats. Aquell coneixement es limitava a una pràctica específica i estava relacionat amb l'execució d'una tasca intel·ligent: diagnòstic, planificació o disseny, per exemple. Els sistemes experts van despertar un gran interès en el món empresarial, des del sector de les assegurances fins al sector aeroespacial, i es van invertir grans quantitats de diners en la seva construcció i el seu manteniment. Però aquells sistemes patien de fragilitat i rigidesa. Tenien enormes dificultats per aprendre coses noves. Com a conseqüència, no eren capaços d'adaptar-se prou ràpidament a canvis en l'entorn i de vegades fracassaven estrepitosament. El camp de la intel·ligència artificial estava en el punt de mira (Lighthill 1973). Van tallar els fons destinats a investigació i les diferents solucions pràctiques i els mètodes que s'estaven ideant es van considerar aliens a la realitat i conceptualment divergents. Aquella ciència s'estava ensorrant. Es van replantejar a fons les seves premisses i els seus objectius inicials. Alguns es van preguntar si raonar era necessari per actuar de manera intel·ligent o si el raonament necessitava una representació simbòlica (Brooks 1991). La lògica era realment necessària per raonar i pensar o amb la mera imitació de la connectivitat de teixits neuronals n'hi hauria prou per fer tasques intel·ligents? I per cert, què era una tasca intel·ligent? Quin era el concepte comú entre, diguem-ne, el processament del llenguatge natural i la robòtica? N'hi havia algun?

La intel·ligència artificial va trigar una mica a reorientar-se. La creació d'un concepte comú que definís l'objecte del seu interès va ser determinant en el seu lent renaixement. El concepte d'agent va reorganitzar els interessos i els objectius de les seves diferents subdivisions (Russell i Norvig 2003). La intel·ligència artificial consistia bàsicament a construir agents intel·ligents que resolguessin problemes sols o de manera conjunta. Per aconseguir-ho, els agents tenien objectius i aplicaven tots els seus coneixements o els coneixements que poguessin obtenir de l'entorn en què operaven. Els agents tenien "agència".

Aquesta agència s'interpretava principalment de dues maneres. Una era una capacitat supeditada a un altre agent, que podia ser un agent humà. És a dir, un agent intel·ligent actuava per compte d'un altre agent (possiblement humà). Per aconseguir-ho tenia una àmplia gamma de maneres alternatives d'assolir els seus objectius per compte d'aquest altre agent (és a dir, els objectius del primer agent eren una traducció dels objectius del segon agent). Segons una definició més forta, agència indicava l'autonomia d'un agent per perseguir els seus propis objectius. Els agents podien "cobrar vida" quan es donés una certa configuració de l'entorn, un estat determinat que els dugués a actuar. Farien el que poguessin (fins i tot aprendre) per assolir aquell objectiu en aquella configuració del món. Els agents podien ser totalment de software o totalment de hardware, però majoritàriament eren una barreja de tots dos. Els sistemes multiagents eren expansions col·lectives d'intel·ligència del paradigma de l'agent (Russell 2003). En qualsevol cas, en la IA moderna un agent és un tipus especial d'agent racional (Simon 1961) que actua en un entorn i persegueix els seus objectius. L'agent coneix l'estat de l'entorn gràcies a sensors i actua en l'esmentat entorn utilitzant efectors. L'acció és el resultat de la decisió a què s'arriba mobilitzant el conjunt de creences de l'agent que relacionen la informació coneguda de l'entorn amb el coneixement. En algunes definicions d'agents intel·ligents, els desitjos representen la situació o les situacions que l'agent voldria aconseguir per arribar als seus objectius. Les *intencions* representen l'estat deliberatiu de l'agent, el que l'agent ha decidit fer.

El concepte de l'agent aclaria la situació de cada subdisciplina de la IA perquè se situava segons la visió que tenia de si mateix. Com a conseqüència, es podia interpretar la robòtica com la creació d'agents de hardware que haurien d'interactuar físicament amb un entorn físic. De manera anàloga, es podia crear un agent de software que actués en un entorn virtual. La disciplina havia trobat finalment una manera comuna de representar i comparar els diferents tipus de sistemes d'IA: la qualitat de la seva agència.

Agents, dades i aprenentatge

Un dels principals canvis que s'han produït en la construcció de sistemes que utilitzen agents intel·ligents des d'aquesta refundació de la intel·ligència artificial és la importància que han adquirit els mètodes d'aprenentatge automàtic. Els sistemes intel·ligents són cada cop més el resultat d'un procés que comença amb dades que semblen ser rellevants per a una tasca determinada. Les dades s'utilitzen com a matèria primera per a un sistema d'aprenentatge automàtic que n'extreu coneixement que es pot aplicar. Per exemple, els sistemes d'aprenentatge automàtic poden ajudar a identificar patrons rellevants en l'àmbit d'una aplicació. Amb un algoritme d'aprenentatge automàtic es crea un model que, quan es crea el sistema posteriorment, li permet decidir si una determinada seqüència d'informació és representativa o no d'un determinat patró. Aquesta és la base dels sistemes de classificació que s'utilitzen habitualment en moltes aplicacions, subsistemes i sistemes pròpiament dits. Per exemple, per identificar cares, classificar situacions, objectes, persones, etc. En general, els mètodes d'aprenentatge automàtic es poden dividir a grans trets en mètodes que proporcionen un criteri per agrupar coses en àmbits desconeguts, mètodes per classificar objectes en àmbits en què sabem quines són les principals classes i, per acabar, mètodes que ajuden un agent a aprendre per si mateix com pot millorar el seu rendiment, és a dir, a fer millor una tasca determinada. Aquests tres tipus d'aprenentatge se solen denominar aprenentatge no supervisat, aprenentatge supervisat i aprenentatge per reforç (Alpaydin 2017). Hi ha altres mètodes que no encaixen tan bé en aquesta classificació, com ara l'aprenentatge per analogia, l'aprenentatge metafòric i l'aprenentatge transferible. Un altre sistema de classificació de l'aprenentatge automàtic dividiria els mètodes d'aprenentatge automàtic en analítico-descriptius o generatius. És a dir, mètodes que ajuden a crear un model per comprendre un àmbit i extreure'n regles de decisió per a la posterior creació del sistema intel·ligent basat en aquest model o mètodes que aprenen de les dades i creen noves coses com ara les xarxes generatives antagòniques (GAN, per les seves sigles en anglès) (Goodfellow et al. 2014).

Fem ara un alto en el camí per veure un exemple que expliqui breument les principals diferències entre els anteriors sistemes d'IA (basats en el coneixement o basats en experts) i els nous sistemes d'aprenentatge automàtic (basats en el Big Data i en el reconeixement de patrons). Imaginem que volem desenvolupar un sistema que sigui capaç de reconèixer un gos en una imatge. Els anteriors sistemes s'entrenarien per reconèixer les característiques d'un gos (forma bàsica, orelles, ulls, pell, nas, etc.) que s'haurien definit d'alguna manera amb anterioritat. Podríem dir, per exemple, que un gos té entre zero i quatre potes. Té entre zero i dues orelles, que poden ser petites o grosses. Però en realitat tot això no va funcionar mai. El món és senzillament massa complex per descriure'l per endavant. Els nous sistemes funcionen al revés: no entrenes el sistema ensenyant-li com és un gos a través d'abstraccions, sinó proporcionant-li milions d'imatges de gossos codificades en una sèrie de números. En el millor dels casos, el sistema aprèn com és un gos reconeixent patrons numèrics que contenen aquestes imatges. Quan mostres al sistema una nova imatge d'un animal pot determinar amb un cert grau de precisió si aquella imatge conté un gos o no. És important assenyalar que els sistemes d'aprenentatge automàtic actuals, encara que siguin capaços d'identificar un gos, mai tenen realment una comprensió semàntica del que és un gos.

En aquest mosaic de mètodes i tècniques, els models basats en xarxes neuronals han conreat cada cop més èxits en els últims quinze anys i la seva aplicació s'ha generalitzat. Són la base dels anomenats mètodes de Deep Learning (Goodfellow et al. 2016).

Actualment la majoria dels sistemes d'IA que s'estan desenvolupant utilitzen profusament les tècniques d'aprenentatge automàtic. És a dir, els actuals agents d'IA tendeixen a ser agents que aprenen amb més possibilitats de tenir autonomia que els anteriors sistemes d'IA, ja que l'aprenentatge automàtic es pot posar en funcionament sempre que els agents decideixin que necessiten adaptar-se millor a les noves situacions. Per construir-los, calen volums de dades més grans, més extensos i més diversos.

Això no vol dir que en tot el procés de crear aquests agents intel·ligents no hi hagi elements humans (no totalment autònoms). Més aviat, si volem que un sistema reconegui gossos autònomament necessitem que els humans diguin al sistema: això és un gos. També necessitem molt de coneixement per utilitzar el mètode d'aprenentatge automàtic adequat per a un determinat àmbit d'aplicació, per provar els models resultants i per ajustar-los abans de posar en funcionament l'agent intel·ligent en el seu entorn operatiu.

3
LA IA EN EL DISSENY

El fet que cada cop més sistemes incloguin algun tipus d'agent intel·ligent en la seva construcció ha obert una àmplia gamma d'interseccions entre objectius, tasques i mètodes de disseny i la intel·ligència artificial. A continuació destaquem algunes de les àrees en què la IA s'aplica al disseny.

Els materials es poden considerar una cosa passiva i una dada coneguda o, al contrari, una cosa que es pot dissenyar perquè exposi un determinat conjunt de propietats; potencialment, els materials es poden

considerar una cosa que pot aconseguir que aquestes propietats evolucionin amb el pas del temps com a resposta a un entorn. L'ús de la intel·ligència artificial sembla especialment indicat per analitzar l'ampli espectre de les maneres de combinar la matèria existent o de modificar els materials existents. Buscar diferents combinacions i operacions per modificar elements individuals o les seves combinacions és característic d'una concepció molt antiga de la IA, que concep el comportament intel·ligent com l'anàlisi intel·ligent d'un espai de recerca realitzat utilitzant l'heurística (Russell i Norvig 2003). En els anys vuitanta ja vam trobar una part d'aquesta visió de l'aplicació de la IA en la creació de nous materials.

Més recentment, també hem trobat l'ús de la "segona onada" de sistemes d'intel·ligència artificial en la creació de nous materials. És a saber, el típic cicle de recopilació de dades, aplicació d'aprenentatge automàtic, provació d'un model i aplicació d'aquest a la tasca que ens ocupa. En aquest cas, s'utilitzen dipòsits i propietats de materials per trobar un conjunt de combinacions i procediments que permetin obtenir el rendiment desitjat d'un substrat material que pot ser un compost o un canvi més fonamental com, per exemple, la pressió, la resposta termal o l'elasticitat. A un altre nivell, es poden analitzar i generar altres propietats de materials utilitzant mètodes d'IA com la textura d'una superfície i també per incloure programabilitat en els mateixos materials (Ion et al. 2017).

La IA s'està utilitzant en el disseny de productes ampliant els seus procediments a nivell del producte. Per exemple, l'obra de Troy Nachtigall estudia la connexió entre les dades i les materialitats (Nachtigall et al. 2019). Utilitzant dades sobre materials, propietats, comportament de l'usuari i altres fonts, s'utilitza un sistema que aprèn de la interconnexió de tots per obtenir un producte ultrapersonalitzat. Els models que se segueixen en l'àmbit del disseny també podrien ser el material d'un sistema d'aprenentatge (Tucker 2016).

L'experiència d'usuari (UX) i la interacció persona-ordinador (HCI) són probablement les subdisciplines que han generat més debat i ús de la IA i de l'aprenentatge automàtic. Hi ha molts casos de sistemes que han utilitzat un tipus o un altre d'IA i d'aprenentatge automàtic per millorar alguns aspectes de la interacció i, concretament, de la interacció persona-màquina. Les àrees dels assistents personals i del disseny d'interfícies són molt actives en aquest sentit. La conceptualització clàssica gira entorn de premisses sobre la funció de les tasques cognitives dels usuaris, dels models simbòlics que, curiosament, s'assemblen a la IA "clàssica" o de "primera onada" (fins als noranta) (Suchman 2007). No obstant això, el comportament dels sistemes d'aprenentatge automàtic respon a complexos models numèrics del món i no a descripcions simbòliques i fa de mitjancer entre la interacció dels usuaris i el món (Blackwell 2015).

La introducció de la IA i de l'aprenentatge automàtic en aquests àmbits també ha obert un interessant debat sobre els mètodes que s'haurien d'utilitzar en la investigació de disseny. La resposta prové de l'adaptació de les tendències actuals, com el disseny centrat en l'usuari (Google Design), a altres propostes més innovadores que reconeixen l'agència dels nous materials i les diferents funcions que tenen en relació amb els usuaris humans, com la integració en comptes d'interacció (Farooq 2106). En alguns casos, la capacitat de predicció dels agents intel·ligents per adaptar-se als usuaris/agents humans (o no humans) ha dut a proposar que aquesta interacció s'interpretés com una variant d'un "disseny anticipatori" general (van Bodegraven 2017).

Col·locar els usuaris humans i els agents artificials al mateix nivell, o gairebé, quant a resultats de l'agència duu a interpretar i dissenyar els sistemes, els serveis i els productes "des del punt de vista de l'agent" (Cruickshank 2017). Això requereix nous mètodes de disseny. Per exemple, "Thing ethnography" (Giaccardi, Cila, Speed i Caldwell 2016a, 2016b) i la investigació del disseny animista (Marenko i van Allen 2016) es podrien considerar aproximacions que adopten un punt de vista no antropocèntric (DiSalvo 2011). Intenten trobar mètodes factibles dins d'un plantejament posthumà del disseny.

4
DESAFIAMENTS

La superposició d'IA, aprenentatge automàtic, Big Data i disseny ha obert noves possibilitats per al disseny. També ha revelat nous problemes que plantegen diferents aspectes del disseny. Sense ànim de ser exhaustius, a continuació analitzem diversos desafiaments.

Assumir les agències artificials

Anar més enllà de la idea que tots els usuaris són necessàriament humans i que el nivell d'agència de les coses dissenyades és molt més elevat que abans posa en dubte alguns dels plantejaments més habituals del disseny. Segurament el disseny centrat en l'usuari és la metodologia més allunyada d'aquesta concepció. No obstant això, la utilitzen els equips que creen els sistemes d'IA (Google Design). Alguns investigadors de disseny comencen a partir d'aquí i reconeixen la naturalesa dinàmica i evolutiva dels sistemes dissenyats amb IA (o per a la IA). Concretament, l'aprenentatge crea noves relacions de coadaptació entre usuaris i sistemes (Leahu 2016). Reconèixer la relació vigent i canviant entre les capacitats dels agents humans i artificials és la base d'enfocaments com el del correndiment (Kuije i Giaccardi 2018), que qüestiona la distribució habitual de l'agència entre agents humans i artificials que formen part del sistema. De fet, tots dos aprenen i canvien perquè funcionen conjuntament i evolucionen conjuntament. Tal com ja hem dit, altres plantejaments com el disseny animista també podrien compartir aquesta idea de l'evolució mútua per mitjà de l'aprenentatge en poblacions intel·ligents mixtes d'agents (humans o artificials) (van Allen 2013).

Connectar la IA, l'aprenentatge automàtic, les dades i el disseny

Tal com s'ha descrit anteriorment, les dades, l'aprenentatge automàtic i els agents d'intel·ligència artificial actualment tenen una relació més estreta que mai. Això imposa nous requisits als professionals del disseny. Cal disposar de les dades adequades per als objectius que preveu el sistema que s'està dissenyant. Aquestes dades, d'alguna manera, haurien de garantir que el model resultant és de qualitat i factible. Tanmateix, els actuals mètodes d'aprenentatge automàtic en particular i la intel·ligència artificial en general avaluen els models des del punt de vista de l'anàlisi científica de dades i no es poden traduir directament en altres criteris d'avaluació que s'utilitzen en disseny i que estan relacionats amb aspectes més qualitatius com, per exemple, la qualitat de l'experiència d'usuari, per dir-ne un (Dove 2017; Kuniavski 2017). Cal disposar de les dades adequades, del mètode d'aprenentatge automàtic adequat i del model resultant adequat per a la tasca que ens ocupa, i continua sent difícil preveure l'impacte que tot això tindrà en els atributs del disseny. Els estudis de l'ús real de dades per a la IA i per al disseny mostren una escissió entre la pràctica habitual en enginyeria de dades i l'aprenentatge automàtic respecte del disseny (Yang 2017). De moment, sembla que l'única manera d'avançar és creant equips multidisciplinaris (Girardin i Lathia 2017).

Posar en pràctica noves interpretacions de conceptes consolidats de disseny

Quan treballem amb sistemes formats per una o més agències artificials, sorgeixen noves situacions que els usuaris humans han d'afrontar i que requereixen una ampliació dels actuals conceptes del disseny. Les noves exigències sotmeten la usabilitat a més pressió. Per exemple, comprendre el comportament d'un agent intel·ligent artificial amb què interactua un usuari, i concretament un agent que aprèn, afegeix pressió a l'enfocament actual basat en plans i cognició (Suchman 2007; Van Allen 2017). Comprendre per què l'agent pren la decisió d'actuar d'una certa manera implica nous aspectes (Seshadri 2017; Huang 2017). Davant d'aquestes preguntes, les respostes que provenen de l'àmbit tècnic (tant si venen de l'aprenentatge automàtic com si venen de la IA o dels sistemes de dades) són complexes i difícils de comunicar a l'usuari de manera efectiva, senzilla i discreta. Els mètodes es divideixen en mètodes que recorren a la transparència (Kizilcec 2016; DTL 2018; Sangüesa 2018) o a l'explicació (Gunning 2018). Aquesta manca de comprensió crea experiències d'usuari de menys qualitat (Springer 2017; Kuniavsky, 2017) i provoca una manca de confiança de l'usuari cap al sistema intel·ligent. No obstant, això és especialment difícil en sistemes derivats d'una aplicació de models neurals, en particular de Deep Learning (Lei 2016). Això minva la confiança en vista de les conegudes implicacions problemàtiques de la IA, l'aprenentatge automàtic i el Big Data: la transmissió i amplificació del biaix (Eubanks 2018; Crawford 2016) i el reforç de la discriminació (Eubanks 2018; Sweeney 2013; FATML 2017, 2018). Si bé actualment continua obert el debat de la creació de sistemes d'IA fiables, aquest enllaça sens cap dubte amb els valors i les pràctiques del disseny (Floridi 2019; IEEE 2018). El disseny ètic dels sistemes intel·ligents requereix combinar aquests aspectes amb responsabilitat (Baylé 2019) i habilitat ètica (Casacuberta i Guersenzvaig 2018). Per acabar, s'hauria d'interpretar l'ètica per a agents intel·ligents –sistemes tecnològics basats en una infraestructura que és mediambientalment exigent– de manera oberta, incloent en la seva consideració no només els humans, sinó també altres entitats. Dit d'una altra manera, hi ha també una relació amb els esforços per dissenyar de manera sostenible en aquest àmbit (DiSalvo 2010). Aquests esforços de sostenibilitat s'han d'augmentar per reduir l'impacte ambiental que genera el desenvolupament dels sistemes d'IA. Hi ha investigacions recents que mostren que la petjada de carboni generada per la formació de models habituals i grans d'IA és gairebé cinc vegades superior a la d'un cotxe americà mitjà, inclosos el mateix cotxe i el combustible que consumeix (Hao 2019). Les sèries numèriques presents en els patrons d'aprenentatge potser serien més pròpies d'un món platònic no afectat per les preocupacions de sostenibilitat, però el hardware que fa possible l'aprenentatge automàtic té efectes ambientals molt reals. Desenvolupar un hardware més sostenible és un imperatiu ètic.

Qualsevol disseny té una dimensió normativa. Des de ressalts que ens obliguen a reduir la velocitat mentre conduïm fins a forquilles o culleres que indiquen de manera aproximada la quantitat de menjar que ens posem a la boca. Els artefactes dissenyats contenen i transmeten normes i principis. En paraules del filòsof de la tecnologia Carl Mitcham (1995): "Diferents dissenys impliquen (de manera implícita o explícita) diferents pressupòsits sociopolítics i diferents visions de la vida. El mateix disseny constitueix una nova manera de dur, o una entrada, a diferents mons de la vida tecnològics". Tot i que el dissenyador no determina mai completament l'ús real dels artefactes, les seves decisions poden tenir i tenen importants conseqüències ètiques. Només pel fet de sentir-se interpel·lats per "com haurien de ser les coses", segons la famosa expressió de Herbert Simon (1961), els dissenyadors assumeixen una tasca èticament exigent per a la qual han d'estar preparats i de la qual han de ser conscients. Això és especialment important quan els dissenyadors no són les úniques agències implicades en les accions que tenen importància des del punt de vista de l'ètica.

5
AQUESTA PUBLICACIÓ

Temes de Disseny va a considerar que havia de convidar a la reflexió sobre aquest tema. Vam fer una crida a investigadors i professionals del disseny perquè compartissin els seus estudis i les seves troballes en intel·ligència artificial, aprenentatge automàtic, Big Data, materials intel·ligents i altres tecnologies emergents i ens plau compartir els assajos que han estat seleccionats després d'una exhaustiva anàlisi realitzada per experts.

Metodologia de disseny de paisatge: crear patrons utilitzant autòmats cel·lulars, de Sergi Abellán, Marcel Bilurbina, i Marilena Christodoulou, mostra com s'ha d'aplicar un constructe artificial consolidat en computació, l'autòmat cel·lular, en el disseny de paisatges. Els autòmats cel·lulars es poden considerar agrupacions d'agents amb regles simples que actuen com el seu "coneixement". Tanmateix, les agrupacions d'aquests autòmats solen tenir comportaments emergents molt complexos que repliquen els comportaments de les colònies d'organismes vius. La comparació de l'evolució de les contraparts artificial i natural presenta un plantejament intel·ligent generatiu en l'àmbit del disseny de paisatge, però també apunta a una aplicabilitat més general.

Disseny d'eines predictives per obtenir funcions personalitzades en peces esportives de punt, de Martijn ten Bhömer, Hai-Ning Liang, Difeng Yu, Yuanjin Liu, Yifan Zhang, Eva de Laat i Carola Leegwater, és una anàlisi de les possibilitats de la indústria 4.0 i de l'impacte que té en el procés del disseny creatiu. La combinació de dades, aprenentatge automàtic i simulació obre possibilitats d'extraordinaris nivells de personalització en la fabricació tèxtil i en el teixit de punt en particular. Aquest article és un estudi de cas en aquest àmbit que analitza de quina manera el software predictiu inspirat en l'aprenentatge automàtic pot ajudar en els processos. Els autors estudien a fons les característiques d'aquest software que pot ser una ajuda per a la creativitat dels dissenyadors, i se centren especialment en el disseny d'interaccions i interfícies.

Crear soft wearables amb i mitjançant tecnologies digitals, Bruna Goveia, Kristina Andersen i Oscar Tomico, examina els últims set anys de funcionament del Wearable Senses Lab. L'estudi se centra en una visió dels wearables intel·ligents que els transforma en ens totalment programables. Els autors amplien l'àmbit de la seva reflexió a tot l'ecosistema que envolta els wearables, i tenen en compte els diferents nivells de personalització, els processos de fabricació i la relació amb serveis associats, usuaris i altres actors. Reserven un interès especial al paper que les dades i la computació tenen en tots els processos.

Practicar moda amb l'Antropocè, de Patricia Wu, proposa la pràctica del disseny des d'una perspectiva especulativa en què l'home deixa de ser el centre, un fenomen característic de l'Antropocè. L'autor analitza les conseqüències d'aquest canvi d'agència en què el món està poblat per entitats no humanes i altament autònomes, els Odradeks. Descrivint el procés pràctic que proposa l'autor s'explicita la possibilitat d'arribar a una pràctica de disseny de moda que encarna plenament una visió ecològica i no centrada en l'home.

La codificació biològica del disseny i les premisses per a una nova generació de productes 'vius': l'exemple de Sinapsi, de Sabrina Lucibello i Carmen Rotondi, descriu el procés mitjançant el qual van crear un dispositiu intel·ligent per ajudar les persones invidents a moure's per diferents entorns. Marcant-se com a objectiu dissenyar un món ple d'intel·ligència artificial, van utilitzar models biològics en el procés de disseny. Segons les autores, d'aquesta manera millora la qualitat de vida de l'usuari d'un sistema intel·ligent i es potencien les qualitats humanes. El fet que provessin la seva tesi en un entorn en què la codependència dels usuaris humans i els sistemes autònoms artificials és tan elevada augmenta l'interès per les seves troballes, ja que podrien apuntar a una manera d'aplicar altres sistemes intel·ligents en altres àmbits.

El disseny d'òrgans en la Transpecies Society: pràctiques híbrides entre cibernètica i intel·ligència artificial, de Tatiana Afanador i Judit Parés. La Transpecies Society és una associació que considera allò cibernètic i allò artificial la seva matèria principal i l'objecte elemental de la seva pràctica. Els autors, d'una manera molt pràctica, tracten els seus mètodes de treball i els llaços que aquests tenen amb una cultura basada en dades, en una relació crítica amb els postulats de la cibernètica i els sistemes intel·ligents. Analitzen les seves pròpies pràctiques observant, per exemple, què passa amb el disseny i el funcionament –o la vida– dels cíborgs.

Plataforma de disseny de serveis per a persones amb demència: teràpia de records centrada en la persona amb intel·ligència artificial en entorns immersius, de Jinyoung Lee, estudia la interconnexió del disseny i els entorns virtuals per facilitar la teràpia de records en persones amb demència i els seus cuidadors. El sistema se centra en l'emmagatzematge, la categorització i la recuperació de records personals dels pacients. S'inclouen els canvis que ha introduït l'ús dels entorns de realitat virtual, ja que, segons l'autora, poden ampliar les actuals possibilitats de representar i reviure els records. També es tenen en compte les implicacions ètiques d'aquests canvis.

Death Inc., del col·lectiu de disseny Domestic Data Streamers, mostra una peça que es va presentar a l'exposició Design Does. L'exposició en qüestió va tenir lloc al Museu del Disseny de Barcelona i girava a l'entorn de diferents dimensions del disseny contemporani i la seva importància en la societat. Death Inc. era una instal·lació sobre robots assassins que volia plantejar que els sistemes d'armes autònoms no segueixen el principi de discriminació. Es basava en el cas del model SGR-A1 de Samsung, que actualment s'utilitza a la zona desmilitaritzada de la frontera entre Corea del Nord i Corea del Sud. Aquest model, considerat el primer de la seva classe, és capaç de seleccionar objectius humans i disparar-los sense que calgui que una persona autoritzi l'operació.

Marc per aplicar sistemàticament l'ètica humanista quan s'utilitza la IA com a material de disseny, de Richelle Dumond, Kyle Dent i Mike Kuniasky, s'enfronta a la dura tasca de configurar un conjunt sòlid i exhaustiu de directrius i pràctiques perquè els dissenyadors puguin respondre als dilemes ètics amb què es troben quan creen sistemes intel·ligents. Fa una atenció especial als problemes que originen els nous nivells d'autonomia d'aquests sistemes i a la manera de tractar-los perquè es respectin els drets humans i la intimitat personal, es garanteixi la confidencialitat de les dades de caràcter personal i es fomentin la llibertat d'expressió i la igualtat, unes condicions, totes, necessàries per sostenir una ètica humanista.

Ens agradaria agrair a tots els autors les seves aportacions, i als revisors, el seu esforç. Esperem que aquesta publicació ajudi els dissenyadors a posicionar-se respecte de la IA i de les tecnologies emergents i que serveixi perquè les seves investigacions i les seves pràctiques generin els millors resultats possibles.

BIOGRAFIES

Ramon Sangüesa
Technical University of Catalonia

Ramon Sangüesa és professor de la Universitat Politècnica de Catalunya (UPC). Anteriorment havia dirigit l'àrea d'investigació tecnològica de l'escola Elisava de disseny i enginyeria i havia coordinat el Data Transparency Lab, un consorci internacional d'investigació en l'àmbit de la transparència de dades. És el fundador de MyData Hub, que forma part de mydata.org, una xarxa de promoció de la sobirania sobre les dades personals. És investigador associat del Center for Organizational Innovation del Departament de Sociologia de la Universitat de Colúmbia a Nova York.

Ariel Guersenzvaig
Elisava School of Design and Engineering

El Dr. Guersenzvaig és el director del màster en Disseny i Comunicació, un màster oficial que dona accés als estudis de doctorat. També dirigeix un curs professional, el màster en Gestió de Projectes i Disseny per a Internet. És professor universitari i de postgrau a Elisava, i professor convidat de diverses universitats (UPF, UPC). Té un doctorat en Teoria del Disseny per la Universitat de Southampton (Regne Unit), un màster en Ètica per la Universitat de Birmingham (Regne Unit), un diploma d'estudis avançats en Investigació en Disseny per la Universitat de Barcelona (Espanya) i un grau d'Edició per la Universitat de Ciències Aplicades d'Àmsterdam (Països Baixos).

REFERÈNCIES

Veure llistat complet de referències a la pàgina 15.

R. Sangüesa y A. Guersenzvaig

La inteligencia artificial como material de diseño: trabajar con nuevas agencias

Traducción al Castellano

PALABRAS CLAVE

Inteligencia Artificial, Diseño, Materiales, Agencia, Aprendizaje Automático, Ética.

1
INTRODUCCIÓN

Inteligencia artificial (IA) se ha convertido en una expresión de uso doméstico, una expresión en boga en los medios de comunicación, pero también en una industria en sí misma y en un conjunto de prácticas asociadas a un negocio y a unos objetivos profesionales. En cuanto tecnología simbólica general con enormes inversiones a sus espaldas, ha influido en muchos ámbitos de actividad. Los *chatbots* pueblan internet, donde se utilizan con objetivos políticos (Bond et al. 2012) y de marketing (Wilson 2017). Por su parte, los sistemas inteligentes regulan las cadenas logísticas (Kückelhaus y Chung 2018), dirigen la atención de la gente hacia determinados productos (Marinchak et al. 2018), influyen en su próxima compra (Portugal et al. 2017), cambian sus emociones (Booth 2014), interpretan pruebas médicas (Koch et al. 2018), interactúan con usuarios de maneras nuevas (Siddike et al. 2018), deciden quién recibe una beca o un préstamo (Agarwal 2019; Eubanks 2018) o cuándo y cómo librar una guerra (Davis 2018). De uno u otro modo, los sistemas inteligentes clasifican, programan, planifican y crean productos, servicios y experiencias. Toman decisiones continuamente en muchos aspectos de nuestra vida, contribuyendo así a darle forma (Turkle 2006).

Pero no es la primera vez que se ha producido un auge en la aplicación de técnicas de IA en la industria. Durante los años 80 y 90 se produjo una explosión de aplicaciones de esta tecnología que despertaron grandes esperanzas, pero la fragilidad que mostró el uso real de las tecnologías de la época terminó por instaurar el primer "Invierno de la IA" (AI Expert Newsletter; Russell y Norvig 2003). Actualmente, el nivel de digitalización en amplios sectores de actividad y la mejora en muchas técnicas de IA han creado el terreno propicio para un nuevo uso generalizado de la IA. Mientras que las versiones antiguas de IA necesitaban el uso de los denominados sistemas expertos, con un alto coste de desarrollo y un alto nivel de dificultad de entrenamiento, la actual combinación de volúmenes y disponibilidad de datos de muchos tipos de aplicaciones ha facilitado el uso de técnicas de aprendizaje automático, que históricamente era una subdisciplina de la IA. El aprendizaje automático ha acelerado la construcción y el entrenamiento de sistemas de IA, superando la antigua escasez de recursos a partir de los cuales se podía extraer conocimiento y con los cuales los sistemas de IA podían llevar a cabo sus tareas.

En la actualidad existe la idea de que la IA tendrá un papel determinante en nuestras sociedades. Es considerada como un activo estratégico que, por poner un ejemplo, afectará a la competencia internacional (Villani 2018; Webster et al. 2017; NSTC 2016; Hogarth 2018). Se ha atribuido a la IA una gran influencia y se le ha confiado, para bien o para mal, la organización de aspectos que van desde el comportamiento personal hasta las estrategias económicas. Sus conexiones con la economía conductual (Pedersen 2018), la vigilancia digital (Zuboff 2015) y las aplicaciones militares (Krishnan 2009) hacen que la preeminencia de la IA como orquestadora de la vida sea un mal presagio y han dado lugar a una intensa actividad en torno a sus implicaciones éticas (Floridi 2019; Dignum 2018; Casacuberta y Guersenzvaig 2018).

En su acepción de práctica profesional y sector en sí mismo, el diseño también ha experimentado las fricciones y el entusiasmo que despierta la IA. Los profesionales de varias áreas del diseño se han visto envueltos en la IA de varias formas. Quizás los primeros que percibieron el impacto de la IA fueron los profesionales del diseño que trabajaban en el ámbito de la interacción persona-ordenador (HCI, por sus siglas en inglés) y de la experiencia de usuario (UX) (Winograd 2006; Grundin 2009), que históricamente ha sido el terreno en el que lo digital ha tenido mayor impacto en el diseño. Pero las técnicas de IA, y especialmente las técnicas de aprendizaje automático, también han hecho incursiones en otros ámbitos del diseño.

Muchos planteamientos generativos del diseño utilizan profusamente la interconexión entre los datos y los algoritmos, especialmente los algoritmos de aprendizaje automático, para idear y producir nuevos productos y la hacen de manera que el algoritmo inteligente asume un papel que es algo más que el papel pasivo o reactivo de una herramienta en manos de los diseñadores (Koch 2017). Además, trabajar con servicios de recomendación y asistentes inteligentes ha comportado nuevos problemas y objetivos para el diseño. Lentamente, nos dimos cuenta de que, en realidad, algunos diseños se crean no con la inteligencia artificial o por la inteligencia artificial, sino para este tipo de entidades (DiSalvo 2011). Los sistemas de IA, pues, a veces tienen al mismo tiempo la función de herramienta, de colaborador y de usuario.

Nos pareció que esta situación merecía una reflexión desde el punto de vista del diseño y de la ingeniería del diseño, y buscamos un tema común que fuera útil para analizar la relación entre la IA y el diseño así como otras tecnologías emergentes como el Big Data o los materiales inteligentes. La IA es considerada básicamente como una entidad "soft" pero, al igual que pasa con otras áreas de la ciencia de la información y la computación, también tiene un componente material (Dourish 2017, 33-59). Pensamos que ese podía ser un buen punto de partida para contrastar distintas interacciones entre la IA y el diseño y, al mismo tiempo, para empezar a clarificar el estado actual de dichas conexiones. Pero al hacerlo también queríamos destacar que en cuanto material de diseño, la IA tiene características muy peculiares. Es difícil interpretar la IA como un material pasivo y constante. Por el contrario, no solo es maleable sino que tiene un cierto nivel de "vida", o al menos de autonomía, en sí misma. Diseñar con IA es claramente trabajar con un material que evolucionará de maneras que son más variadas y más difíciles de planificar por adelantado, al menos en comparación con otros materiales más tradicionales. Trabajar con madera contrachapada o con aluminio está circunscrito totalmente a las leyes de la física, algo que los convierte en materiales altamente predecibles. Los algoritmos de IA no son ajenos al terreno de la física, por supuesto, pero a causa de su misma naturaleza su comportamiento podría ser muy complicado de vaticinar en detalle. Por el momento, no existe ningún modo directo de analizar un modelo de IA y hacer que sea totalmente predecible.

¿Existe algún concepto común que, siguiendo el mismo razonamiento, nos pueda ayudar a estudiar los distintos tipos de diseño que se están creando en relación a la IA, el aprendizaje automático y los datos? ¿Existe algún concepto que ambas disciplinas utilicen y que nos pueda ayudar a fijar las relaciones que hay entre ellas? Tras un tiempo de reflexión y debate llegamos a la conclusión de que la noción de "agencia" era una buena candidata para ayudarnos a dibujar una conexión práctica con los objetos de la IA, el aprendizaje automático, el Big Data y el diseño. Cuando menos, ciertas similitudes de conceptualización y uso práctico que había en ambas disciplinas eran intrigantes.

La diferencia que la IA puede introducir en general –y en el diseño en particular– podía interpretarse como un salto en el nivel de agencia debido al uso de la IA como material de diseño. En el diseño se ha introducido y utilizado la "agencia" en varias subdisciplinas y prácticas. Un ejemplo es la investigación de diseño, pero también se ha aplicado como guía en la práctica del diseño. La noción de "agencia" arrastra cierta tradición, por ejemplo, de campos como la Actor-Network Theory (Latour 2005; Yaneva 2009). En diseño, se reconoce la agencia de objetos, que puede guardar relación con conceptos como las affordances o puede aplicarse en investigación de usuario. De algún modo, la agencia es un concepto aceptado en la práctica del diseño. Conceptos como los "objetos intencionales" (Rozendaal 2016), por ejemplo, guardan una gran semejanza con los tipos de agencia que se utilizan en IA. Se podría decir que ahora la IA, en efecto, es la disciplina de crear agentes artificiales con sus propios objetivos, deseos e intenciones. Los académicos de la postfenomenología (por ejemplo, Verbeek 2011) articulan una perspectiva filosófica en la que la agencia (moral) se reduce a una cuestión de mediación entre personas y tecnologías en lugar de una cuestión puramente humana.

De hecho, en inteligencia artificial, el concepto de agencia y agencias se ha convertido en un elemento central y es un constructo fundamental para la construcción de sistemas complejos de IA (Wooldridge 2002).

Existen diferencias entre las conceptualizaciones utilizadas en cada disciplina. Para diferenciarlas podríamos retroceder un poco y analizar los orígenes y los objetivos de la IA y por qué y cómo el concepto de agencia actualmente desempeña un papel fundamental en varias disciplinas.

2
LA AGENCIA EN LOS SISTEMAS DE INTELIGENCIA ARTIFICIAL

El inicio de la inteligencia digital como disciplina suele asociarse con la celebración en 1956 de la Escuela de Verano de Dartmouth sobre Inteligencia Artificial (McCarthy et al. 1955). Aquel encuentro reunió a investigadores en matemáticas, tecnología de la información, lógica, computación y psicología. Todos compartían un mismo entusiasmo por las posibilidades que abrían los ordenadores. Los pioneros de Dartmouth previeron que aquellas máquinas podían superar el uso que tenían por aquel entonces. Pensaron que los ordenadores podían ser entidades capaces de hacer mucho más que simplemente procesar números. Pensaron que los ordenadores podrían manipular muchos tipos distintos de representaciones, incluidas representaciones del conocimiento. El pensamiento podía ser, como había dicho ya Hobbes siglos atrás, una mera cuestión de cálculo. Aunque esta visión de cálculo de nuevos tipos de símbolos, y no solo de números, era la dominante en la idea de inteligencia de las máquinas que caracterizó el encuentro de Dartmouth, otra parte del mismo grupo original estaba más interesada en cálculos más simples. Pensaban que llegarían a la inteligencia no con estructuras simbólicas complejas sino replicando la simple combinación de señales de la infraestructura neuronal de los sistemas inteligentes cognitivos. El grupo de Dartmouth emitió una declaración que, más que un programa detallado de investigación científica, tenía la apariencia de una lista de ejemplos de tareas que asociaban al comportamiento inteligente y que pensaban que había que estudiar. Vale la pena echar una ojeada a la lista (McCarthy et al. 1955), pues arroja luz sobre la ambición del planteamiento del grupo de Dartmouth y también sobre los límites de la imaginación de aquel grupo fundacional: teniendo en cuenta la capacidad que tenían los ordenadores en aquella época, había que ser valiente para obtener de aquellas máquinas las tareas extremamente más complejas que necesitarían en el futuro para atribuirles inteligencia. Era un ejercicio de especulación y proyección, quizás inspirado por un clima de investigación empapado de los logros de Alan Turing y otros científicos en los años treinta y que llevaron a Turing a escribir su artículo fundacional "Maquinaria computacional e inteligencia" (Turing 1950). Uno de los participantes en aquel histórico encuentro definió el objetivo global de la incipiente disciplina de la inteligencia artificial con las siguientes palabras:

"Estudiar cómo crear sistemas que se comporten de un modo que, si fueran humanos, se pueda considerar 'inteligente'". (Minsky 1986)

Curiosamente, cada uno de los participantes en aquella naciente disciplina abordaba el objetivo común desde la práctica de la disciplina de la que provenía. Más que una metodología científica –entendida como los métodos de las ciencias duras teóricas o experimentales–, preferían un planteamiento constructivo o, según una terminología posterior, un planteamiento construccionista. Es decir, sabrían cómo alcanzar aquellos objetivos construyendo sistemas que los lograran y aprendiendo del proceso. Estos son básicamente los rasgos de una "cultura del diseño" (Serra 1992), tal como podría aceptar Herbert Simon, uno de los más destacados asistentes a los encuentros de Dartmouth (Simon 1961).

Aquel encuentro dio lugar a muchas actividades y rápidamente empezaron a producirse gran número de éxitos iniciales. A finales de los ochenta y principios de los noventa, el ámbito más prometedor de la IA era el de sistemas expertos. Estos se convirtieron en sistemas de conocimiento, y dieron lugar a una disciplina totalmente nueva: la ingeniería del conocimiento. Los sistemas expertos eran básicamente sistemas de razonamiento que utilizaban una representación del conocimiento acumulado por expertos humanos altamente cualificados. Aquel conocimiento se limitaba a una práctica específica y estaba relacionado con la ejecución de una tarea inteligente: diagnóstico, planificación o diseño, por ejemplo. Los sistemas expertos despertaron un gran interés en el mundo empresarial, desde el sector de seguros hasta el sector aeroespacial, y se invirtieron grandes sumas de dinero en su construcción y mantenimiento. Pero aquellos sistemas adolecían de fragilidad y rigidez. Tenían enormes dificultades por aprender nuevas cosas. Por consiguiente, no eran capaces de adaptarse con suficiente rapidez a cambios en el entorno, y a veces fracasaban estrepitosamente. El campo de la inteligencia artificial estaba en el punto de mira (Lighthill 1973). Cortaron los fondos destinados a investigación y las distintas soluciones prácticas y métodos que se estaban ideando se consideraron ajenos a la realidad y conceptualmente divergentes. Aquella ciencia se estaba desmoronando. Se replantearon a fondo sus premisas y objetivos iniciales. Algunos se preguntaron si razonar era necesario para actuar de manera inteligente o si el razonamiento necesitaba una representación simbólica (Brooks 1991). ¿La lógica era realmente necesaria para razonar y pensar o la mera imitación de la conectividad de tejidos neuronales sería suficiente para realizar tareas inteligentes? Y por cierto, ¿qué era una tarea inteligente? ¿Cuál era el concepto común entre, digamos, el procesamiento del lenguaje natural y la robótica? ¿Había alguno?

La inteligencia artificial tardó un tiempo en reorientarse. La creación de un concepto común que definiera el objeto de su interés fue determinante en su lento renacer. El concepto de agente reorganizó los intereses y los objetivos de sus distintas subdivisiones (Russell y Norvig 2003). La inteligencia artificial consistía básicamente en construir agentes inteligentes que resolvieran problemas solos o conjuntamente. Para lograrlo, los agentes tenían objetivos y aplicaban todos sus conocimientos o los conocimientos que pudieran obtener del entorno en el que operaban. Los agentes tenían "agencia".

Esta agencia se interpretaba principalmente de dos maneras. Una era una capacidad supeditada a otro agente, que podía ser un agente humano. Es decir, un agente inteligente actuaba por cuenta de otro agente (posiblemente humano). Para lograrlo tenía una amplia gama de maneras alternativas de alcanzar sus objetivos por cuenta de este otro agente (es decir, los objetivos del primer agente eran una traducción de los objetivos del segundo agente). Según una definición más fuerte, agencia indicaba la autonomía de un agente por perseguir sus propios objetivos. Los agentes podían "cobrar vida" cuando se diera una cierta configuración del entorno,

un estado determinado que los llevara a actuar. Harían lo que pudieran (incluso aprender) para lograr aquel objetivo en aquella configuración del mundo. Los agentes podían ser totalmente de software o totalmente de hardware, pero mayoritariamente eran una mezcla de ambos. Los sistemas multiagentes eran expansiones colectivas de inteligencia del paradigma del agente (Russell 2003). En cualquier caso, en la IA moderna un agente es un tipo especial de agente racional (Simon 1961) que actúa en un entorno y persigue sus objetivos. El agente conoce el estado del entorno gracias a sensores y actúa en dicho entorno utilizando efectores. La acción es el resultado de la decisión a la que se llega movilizando el conjunto de creencias del agente que relacionan la información conocida del entorno con el conocimiento. En algunas definiciones de agentes inteligentes, los deseos representan la situación o situaciones que el agente querría lograr para llegar a sus objetivos. Las intenciones representan el estado deliberativo del agente, lo que el agente ha decidido hacer.

El concepto del agente aclaraba la situación de cada subdisciplina de la IA porque se situaba en función de la visión que tenía de sí mismo. Por consiguiente, se podía interpretar la robótica como la creación de agentes de hardware que deberían interactuar físicamente con un entorno físico. De manera análoga, se podía crear un agente de software que actuaría en un entorno virtual. La disciplina había encontrado finalmente una manera común de representar y comparar los distintos tipos de sistemas de IA: la calidad de su agencia.

Agentes, datos y aprendizaje

Uno de los principales cambios que se han producido en la construcción de sistemas que utilizan agentes inteligentes desde esta refundación de la inteligencia artificial es la importancia que han adquirido los métodos de aprendizaje automático. Los sistemas inteligentes son cada vez más el resultado de un proceso que empieza con datos que parecen ser relevantes para una determinada tarea. Los datos se utilizan como materia prima para un sistema de aprendizaje automático que extrae conocimiento que se puede aplicar. Por ejemplo, los sistemas de aprendizaje automático pueden ayudar a identificar patrones relevantes en el ámbito de una aplicación. Con un algoritmo de aprendizaje automático se crea un modelo que, cuando posteriormente se crea el sistema, le permite decidir si una determinada secuencia de información es representativa o no de un determinado patrón. Esa es la base de los sistemas de clasificación que se utilizan habitualmente en muchas aplicaciones, subsistemas y sistemas propiamente dichos. Por ejemplo, para identificar caras, clasificar situaciones, objetos, personas, etc. En general, los métodos de aprendizaje automático se pueden dividir a grandes rasgos en métodos que proporcionan un criterio para agrupar cosas en ámbitos desconocidos, métodos para clasificar objetos en ámbitos en los que conocemos cuáles son las principales clases y, por último, métodos que ayudan a un agente a aprender por sí mismo cómo mejorar su rendimiento, es decir, hacer mejor una determinada tarea. Estos tres tipos de aprendizaje suelen denominarse aprendizaje no supervisado, aprendizaje supervisado y aprendizaje por refuerzo (Alpaydin 2017). Existen otros métodos que no encajan tan bien en esta clasificación, como el aprendizaje por analogía, el aprendizaje metafórico y el aprendizaje transferible. Otro sistema de clasificación del aprendizaje automático dividiría los métodos de aprendizaje automático en analítico-descriptivos o generativos. Es decir, métodos que ayudan a crear un modelo para comprender un ámbito y extraer reglas de decisión para la posterior creación del sistema inteligente basado en dicho modelo o métodos que aprenden de los datos y crean nuevas cosas como las redes generativas antagónicas (GAN, por sus siglas en inglés) (Goodfellow et al. 2014).

Hagamos un alto en el camino para ver un ejemplo que explique brevemente las principales diferencias entre los anteriores sistemas de IA (basados en el conocimiento o basados en expertos) y los nuevos sistemas de aprendizaje automático (basados en el Big Data y en el reconocimiento de patrones). Imaginemos que queremos desarrollar un sistema que sea capaz de reconocer un perro en una imagen. Los anteriores sistemas se entrenarían para reconocer las características de un perro (forma básica, orejas, ojos, piel, nariz, etc.) que se habrían definido de alguna manera con anterioridad. Podríamos decir, por ejemplo, que un perro tiene entre cero y cuatro patas. Tiene entre cero y dos orejas, que pueden ser pequeñas o grandes. Pero en realidad todo eso nunca funcionó. El mundo es simplemente demasiado complejo para describirlo por adelantado. Los nuevos sistemas funcionan al revés: entrenas al sistema no enseñándole cómo es un perro a través de abstracciones, sino proporcionándole millones de imágenes de perros codificadas en una serie de números. En el mejor de los casos, el sistema aprende cómo es un perro reconociendo patrones numéricos que contienen esas imágenes. Cuando le muestras al sistema una nueva imagen de un animal, el sistema puede determinar con un cierto grado de precisión si aquella imagen contiene un perro o no. Es importante tener en cuenta que los sistemas de aprendizaje automático actuales, incluso si son capaces de identificar a un perro, en realidad nunca tienen una comprensión semántica de lo que es un perro.

En este mosaico de métodos y técnicas, los modelos basados en redes neuronales han cosechado cada vez más éxitos en los últimos quince años y su aplicación se ha generalizado. Son la base de los llamados métodos de Deep Learning (Goodfellow et al. 2017).

En la actualidad la mayoría de sistemas de IA que se están desarrollando usan profusamente las técnicas de aprendizaje automático. Es decir, los actuales agentes de IA tienden a ser agentes que aprenden con más posibilidades de tener autonomía que los anteriores sistemas de IA ya que el aprendizaje automático puede ponerse en funcionamiento siempre que los agentes decidan que necesitan adaptarse mejor a nuevas situaciones. Para construirlos, hacen falta volúmenes de datos mayores, más extensos y más diversos.

Eso no significa que en todo el proceso de crear estos agentes inteligentes no haya elementos humanos (no totalmente autónomos). Más bien, si queremos que un sistema reconozca autónomamente a perros necesitamos que los humanos le digan al sistema: esto es un perro. También necesitamos mucho conocimiento para utilizar el método de aprendizaje automático adecuado para un determinado ámbito de aplicación, para probar los modelos resultantes y para ajustarlos antes de poner en funcionamiento al agente inteligente en su entorno operativo.

3
LA IA EN EL DISEÑO

El hecho de que cada vez más sistemas incluyan algún tipo de agente inteligente en su construcción ha abierto una amplia gama de intersecciones entre objetivos, tareas y métodos de diseño y la inteligencia artificial. A continuación destacamos algunas de las áreas en las que se aplica la IA al diseño.

Los materiales se pueden considerar algo pasivo y un dato conocido o, por el contrario, algo que se puede diseñar para que exponga un determinado conjunto de propiedades; potencialmente, los materiales se pueden considerar algo que puede lograr que dichas propiedades evolucionen con el paso del tiempo como respuesta a un entorno. El uso de la inteligencia artificial parece especialmente indicado para analizar el amplio espectro de las maneras de combinar la materia existente o de modificar los materiales existentes. Buscar diferentes combinaciones y operaciones para modificar elementos individuales o sus combinaciones es característico de una concepción muy antigua de la IA, que concibe el comportamiento inteligente como el análisis inteligente de un espacio de búsqueda realizado utilizando la heurística (Russell y Norvig 2003). Encontramos parte de esta visión de la aplicación de la IA en la creación de nuevos materiales ya en los años 80.

Más recientemente, también encontramos el uso de la "segunda ola" de sistemas de inteligencia artificial en la creación de nuevos materiales. A saber, el típico ciclo de recopilación de datos, aplicación de aprendizaje automático, prueba de un modelo y aplicación del mismo a la tarea que nos ocupa. En este caso, se utilizan depósitos de procesos y propiedades de materiales para dar con un conjunto de combinaciones y procedimientos que permitan obtener el rendimiento deseado de un sustrato material que puede ser un compuesto o un cambio más fundamental como, por ejemplo, la presión, la respuesta termal o la elasticidad. A un nivel distinto, se pueden analizar y generar otras propiedades de materiales utilizando métodos de IA como la textura de una superficie y también para incluir programabilidad en los mismos materiales (Ion et al. 2017).

La IA se está utilizando en el diseño de productos ampliando sus procedimientos a nivel del producto. Por ejemplo, la obra de Troy Nachtigall estudia la conexión entre los datos y las materialidades (Nachtigall et al. 2019). Utilizando datos sobre materiales, propiedades, comportamiento del usuario y otras fuentes, se utiliza un sistema que aprende de la interconexión de todos ellos para obtener un producto ultrapersonalizado. Los modelos que se siguen en el ámbito del diseño también podrían ser el material de un sistema de aprendizaje (Tucker 2016).

La experiencia de usuario (UX) y la interacción persona-ordenador (HCI) son probablemente las subdisciplinas que han generado más debate y uso de la IA y del aprendizaje automático. Existen muchos casos de sistemas que han utilizado uno u otro tipo de IA y de aprendizaje automático para mejorar ciertos aspectos de la interacción y, concretamente, de la interacción persona-máquina. Las áreas de los asistentes personales y del diseño de interfaces son muy activas en ese sentido. La conceptualización clásica gira en torno a premisas sobre la función de las tareas cognitivas de los usuarios, de los modelos simbólicos que, curiosamente, tienen similitudes con la IA "clásica" o de "primera ola" (hasta los 90) (Suchman 2007). No obstante, el comportamiento de los sistemas de aprendizaje automático responde a complejos modelos numéricos del mundo y no a descripciones simbólicas, mediando entre la interacción de los usuarios y el mundo (Blackwell 2015).

La introducción de la IA y del aprendizaje automático en estos ámbitos también ha abierto un interesante debate sobre los métodos que deberían utilizarse en la investigación de diseño. La respuesta proviene de la adaptación de las tendencias actuales, como el diseño centrado en el usuario (Google Design), a otras propuestas más innovadoras que reconocen la agencia de los nuevos materiales y las distintas funciones que tienen en relación a los usuarios humanos como la integración en lugar de la interacción (Farooq 2106). La capacidad de predicción de los agentes inteligentes para adaptarse a los usuarios/agentes humanos (o no humanos) ha llevado a algunos a proponer que dicha interacción debería interpretarse como una variante de un "diseño anticipatorio" general (van Bodegraven 2017).

Colocar a los usuarios humanos y a los agentes artificiales al mismo nivel, o casi, en cuanto a resultados de la agencia lleva a interpretar y diseñar los sistemas, los servicios y los productos "desde el punto de vista del agente" (Cruickshank 2017). Eso requiere nuevos métodos de diseño. Por ejemplo, "Thing ethnography" (Giaccardi, Cila, Speed y Caldwell 2016a, 2016b) y la investigación del diseño animista (Marenko y van Allen 2016) se podrían considerar aproximaciones que adoptan un punto de vista no antropocéntrico (DiSalvo 2011). Intentan dar con métodos factibles dentro de un planteamiento post-humano del diseño.

4
DESAFÍOS

La superposición de IA, aprendizaje automático, Big Data y diseño ha abierto nuevas posibilidades para el diseño. También ha revelado nuevos problemas que plantean distintos aspectos del diseño. Sin ánimo de ser exhaustivos, analizamos a continuación varios desafíos.

Asumir las agencias artificiales

Ir más allá de la idea de que todos los usuarios son necesariamente humanos y de que el nivel de agencia de las cosas diseñadas es mucho mayor que antes pone en tela de juicio algunos de los planteamientos más habituales del diseño. Seguramente el diseño centrado en el usuario es la metodología más alejada de esta concepción. No obstante, lo utilizan los equipos que crean los sistemas de IA (Google Design). Ciertos investigadores de diseño empiezan a partir de ahí y reconocen la naturaleza dinámica y evolutiva de los sistemas diseñados con IA (o para la IA). Concretamente, el aprendizaje crea nuevas relaciones de co-adaptación entre usuarios y sistemas (Leahu 2016). Reconocer la relación vigente y cambiante entre las capacidades de los agentes humanos y artificiales es la base de enfoques como el del co-rendimiento (Kuije y Giaccardi 2018), que cuestiona la distribución habitual de la agencia entre agentes humanos y artificiales que forman parte del sistema. De hecho, ambos aprenden y cambian porque funcionan conjuntamente y evolucionan conjuntamente. Como hemos dicho antes, otros planteamientos como el diseño animista también podrían compartir esta idea de la evolución mutua mediante el aprendizaje en poblaciones inteligentes mixtas de agentes (humanos o artificiales) (van Allen 2013).

Conectar la IA, el aprendizaje automático, los datos y el diseño

Tal como se ha descrito anteriormente, los datos, el aprendizaje automático y los agentes de inteligencia artificial en la actualidad guardan una relación más estrecha que nunca. Eso impone nuevos requisitos a los profesionales del diseño. Es necesario disponer de los datos adecuados para los objetivos que prevé el sistema que se está diseñando. Esos datos, de algún modo, deberían garantizar que el modelo resultante es de calidad y factible. Sin embargo, los actuales métodos de aprendizaje automático en particular y la inteligencia artificial en general evalúan los modelos desde el punto de vista del análisis científico de datos y no se pueden traducir directamente en otros criterios de evaluación que se utilizan en diseño y que guardan relación con aspectos más cualitativos como, por ejemplo, la calidad de la experiencia de usuario, por nombrar uno (Dove 2017; Kuniavski 2017). Es necesario disponer de los datos adecuados, del método de aprendizaje automático adecuado y del modelo resultante adecuado para la tarea que nos ocupa, y sigue siendo difícil prever el impacto que todo ello tendrá en los atributos del diseño. Los estudios del uso real de datos para la IA y para el diseño muestran una escisión entre la práctica habitual en ingeniería de datos y el aprendizaje automático con respecto al diseño (Yang 2017). Por el momento, parece que la única manera de avanzar es creando equipos multidisciplinarios (Girardin y Lathia 2017).

Poner en práctica nuevas interpretaciones de conceptos afianzados de diseño

Cuando trabajamos con sistemas formados por una o más agencias artificiales, surgen nuevas situaciones que los usuarios humanos deben afrontar y que requieren una ampliación de los actuales conceptos del diseño. Las nuevas exigencias someten la usabilidad a mayor presión. Por ejemplo, comprender el comportamiento de un agente inteligente artificial con el que interactúa un usuario, y concretamente un agente que aprende, añade presión al enfoque actual basado en planes y cognición (Suchman 2007; Van Allen 2017). Comprender por qué el agente toma la decisión de actuar de un cierto modo implica nuevos aspectos (Seshadri 2017; Huang 2017). Ante estas preguntas, las respuestas que vienen del ámbito técnico (tanto si vienen del aprendizaje automático, de la IA o de los sistemas de datos) son complejas y difíciles de comunicar al usuario de manera efectiva, sencilla y discreta. Los métodos actuales se dividen en métodos que recurren a la transparencia (Kizilcec 2016; DTL 2018; Sangüesa 2018) o a la explicación (Gunning 2018). Esta falta de comprensión crea experiencias de usuario de menor calidad (Springer 2017; Kuniavsky 2017) y una falta de confianza del usuario hacia el sistema inteligente. Sin embargo, esto es especialmente difícil en sistemas derivados de una aplicación de modelos neuronales, en particular Deep Learning (Lei 2016). Eso merma la confianza en vista de las conocidas implicaciones problemáticas de la IA, el aprendizaje automático y el Big Data: la transmisión y amplificación del sesgo (Eubanks 2018; Crawford 2016) y el refuerzo de la discriminación (Eubanks 2018; Sweeney 2013; FATML 2017, 2018). Si bien actualmente sigue abierto el debate de la creación de sistemas de IA fiables, este enlaza sin duda alguna con los valores y las prácticas del diseño (Floridi 2019; IEEE 2018). El diseño ético de los sistemas inteligentes requiere combinar estos aspectos con responsabilidad (Baylé 2019) y habilidad ética (Casacuberta y Guersenzvaig 2018). Por último, debería interpretarse la ética para agentes inteligentes –sistemas tecnológicos basados en una infraestructura que es medioambientalmente exigente– de manera abierta, incluyendo en su consideración no solo a los humanos sino también a otras entidades. Dicho de otro modo, existe también una relación con los esfuerzos por diseñar de manera sostenible en este ámbito (DiSalvo 2010). Estos esfuerzos de sostenibilidad deben ampliarse para reducir el impacto ambiental que genera el desarrollo de los sistemas de IA. Hay investigaciones recientes que muestran que la huella de carbono generada por el entrenamiento de los tipos habituales de modelos grandes de IA es casi cinco veces mayor que la de un automóvil estadounidense típico, incluido el propio automóvil y el combustible que consume (Hao 2019). Las series numéricas presentes en los patrones de aprendizaje quizás se consideren parte de un reino platónico abstracto que no se ve afectado por cuestiones de sostenibilidad, pero el hardware que hace posible el aprendizaje automático tiene efectos ambientales muy reales. Desarrollar un hardware más sostenible es un imperativo ético.

Todo diseño tiene una dimensión normativa. Desde badenes que nos obligan a reducir la velocidad al conducir hasta tenedores o cucharas que indican de manera aproximada la cantidad de comida que nos ponemos en la boca. Los artefactos diseñados contienen y transmiten normas y principios. En palabras del filósofo de la tecnología Carl Mitcham (1995): "Distintos diseños implican (de manera implícita o explícita) distintos presupuestos sociopolíticos y distintas visiones de la vida. El mismo diseño constituye una nueva manera de llevar, o una entrada, a distintos mundos de la vida tecnológicos". Aunque el diseñador nunca determina por completo el uso real de los artefactos, las decisiones de estos pueden tener y tienen importantes consecuencias éticas. Por el mero hecho de sentirse interpelados por "cómo

deberían ser las cosas", según la famosa expresión de Herbert Simon (1961), los diseñadores asumen una tarea éticamente exigente para la que deben estar preparados y de la que deben ser conscientes. Eso es especialmente importante cuando los diseñadores no son las únicas agencias implicadas en las acciones que revisten importancia desde el punto de vista de la ética.

5
ESTA PUBLICACIÓN

Temes de Disseny consideró que había que invitar a la reflexión alrededor de este tema. Hicimos un llamamiento a investigadores y profesionales del diseño para que compartieran sus estudios y sus hallazgos en inteligencia artificial, aprendizaje automático, Big Data, materiales inteligentes y otras tecnologías emergentes. Nos complace compartir los ensayos que han sido seleccionados tras un exhaustivo análisis realizado por expertos.

Metodología de diseño de paisaje: crear patrones utilizando autómatas celulares, de Sergi Abellán, Marcel Bilurbina y Marilena Christodoulou, muestra cómo aplicar un constructo artificial consolidado en computación, el autómata celular, en el diseño de paisajes. Los autómatas celulares pueden considerarse agrupaciones de agentes con reglas simples que actúan como su "conocimiento". Sin embargo, las agrupaciones de dichos autómatas suelen tener comportamientos emergentes muy complejos que replican los comportamientos de las colonias de organismos vivos. La comparación de la evolución de las contrapartes artificial y natural presenta un planteamiento inteligente generativo en el ámbito del diseño de paisaje pero también apunta a una aplicabilidad más general.

Diseño de herramientas predictivas para obtener funciones personalizadas en prendas deportivas de punto, de Martijn ten Bhömer, Hai-Ning Liang Difeng Yu, Yuanjin Liu, Yifan Zhang, Eva de Laat y Carola Leegwater, es un análisis de las posibilidades de la industria 4.0 y del impacto que tiene en el proceso del diseño creativo. La combinación de datos, aprendizaje automático y simulación abre posibilidades de extraordinarios niveles de personalización en la fabricación textil y en el tejido de punto en particular. Este artículo es un estudio de caso en este ámbito que analiza cómo puede el software predictivo inspirado en el aprendizaje automático ayudar en los procesos. Los autores estudian a fondo las características de este software que puede ser una ayuda para la creatividad de los diseñadores, y se centran especialmente en el diseño de interacciones e interfaces.

Crear soft wearables con y mediante tecnologías digitales, de Bruna Goveia, Kristina Andersen y Oscar Tomico, examina los últimos siete años de funcionamiento del Wearable Senses Lab. El estudio se centra en una visión de los wearables inteligentes que los transforma en entes totalmente programables. Los autores amplían el ámbito de su reflexión a todo el ecosistema que rodea los wearables, y tienen en cuenta los distintos niveles de personalización, los procesos de fabricación y la relación con servicios asociados, usuarios y otros actores. Reservan un interés especial al papel que tienen los datos y la computación en todos los procesos.

Practicar moda con el Antropoceno, de Patricia Wu, propone la práctica del diseño desde una perspectiva especulativa en la que el hombre deja de ser el centro, un fenómeno característico del Antropoceno. El autor analiza las consecuencias de este cambio de agencia en el que el mundo está poblado por entidades no humanas y altamente autónomas, los Odradeks. Describiendo el proceso práctico que propone el autor se explicita la posibilidad de llegar a una práctica de diseño de moda que encarna plenamente una visión ecológica y no centrada en el hombre.

La codificación biológica del diseño y las premisas para una nueva generación de productos 'vivos': el ejemplo de Sinapsi, de Sabrina Lucibello y Carmen Rotondi, describe el proceso mediante el que crearon un dispositivo inteligente para ayudar a las personas invidentes a moverse por diferentes entornos. Marcándose como objetivo diseñar un mundo lleno de inteligencia artificial, se sirvieron de modelos biológicos en el proceso de diseño. Según las autoras, de ese modo mejora la calidad de vida del usuario de un sistema inteligente y se potencian las cualidades humanas. El hecho de que probaran su tesis en un entorno en el que la co-dependencia de los usuarios humanos y los sistemas autónomos artificiales es tan elevada aumenta el interés por sus hallazgos, ya que podrían apuntar a una manera de aplicar otros sistemas inteligentes en otros ámbitos.

El diseño de órganos en la Transpecies Society: prácticas híbridas entre cibernética e inteligencia artificial, de Tatiana Afanador y Judit Parés. La Transpecies Society es una asociación que considera lo cibernético y lo artificial como su material principal y como el objeto elemental de su práctica. Los autores abordan, de manera muy práctica, sus métodos de trabajo y los lazos que estos tienen con una cultura basada en datos, en una relación crítica con los postulados de la cibernética y los sistemas inteligentes. Analizan sus propias prácticas observando, por ejemplo, qué sucede con el diseño y el funcionamiento –o la vida– de los cíborgs.

Plataforma de diseño de servicios para personas con demencia: terapia de recuerdos centrada en la persona con inteligencia artificial en entornos inmersivos, de Jinyoung Lee, estudia la interconexión del diseño y los entornos virtuales para facilitar la terapia de recuerdos en personas con demencia y sus cuidadores. El sistema se centra en el almacenamiento, la categorización y la recuperación de recuerdos personales de los pacientes. Se incluyen los cambios que ha introducido el uso de los entornos de realidad virtual ya que, en opinión de la autora, pueden ampliar las actuales posibilidades de representar y revivir los recuerdos. También se tienen en cuenta las implicaciones éticas de dichos cambios.

Death Inc., del colectivo de diseño Domestic Data Streamers, muestra una pieza que se presentó en la exposición "Design Does". La exposición en cuestión tuvo lugar en el Museo del Diseño de Barcelona y giraba en torno a distintas dimensiones del diseño contemporáneo y su importancia en la sociedad. Death Inc. era una instalación sobre robots asesinos que quería plantear que los sistemas de armas autónomos no siguen el principio de discriminación. Se basaba en el caso del modelo SGR-A1 de Samsung, que actualmente se utiliza en la zona desmilitarizada de la frontera entre Corea del Norte y Corea del Sur. Este modelo, considerado el primero de su clase, es capaz de seleccionar objetivos humanos y dispararles sin que sea necesaria la autorización de una persona para la operación.

Marco para aplicar sistemáticamente la ética humanista cuando se utiliza la IA como material de diseño, de Richelle Dumond, Kyle Dent y Mike Kuniasky, se enfrenta a la dura tarea de configurar un conjunto sólido y exhaustivo de directrices y prácticas para que los diseñadores puedan responder a los dilemas éticos con los que se encuentran cuando crean sistemas inteligentes. Presta una atención especial a los problemas que originan los nuevos niveles de autonomía de estos sistemas y a cómo abordarlos de manera que se respeten los derechos humanos y la intimidad personal, se garantice la confidencialidad de los datos de carácter personal y se fomenten la libertad de expresión y la igualdad, condiciones, todas ellas, necesarias para sostener una ética humanista.

Quisiéramos agradecer a todos los autores por sus aportaciones, y a los revisores por su esfuerzo. Esperamos que esta publicación ayude a los diseñadores a posicionarse respecto de la IA y de las tecnologías emergentes y que sirva para que sus investigaciones y sus prácticas generen los mejores resultados posibles.

BIOGRAFÍAS

Ramon Sangüesa
Technical University of Catalonia

Ramon Sangüesa es profesor de la Universitat Politècnica de Catalunya (UPC). Anteriormente había dirigido el área de investigación tecnológica de la escuela Elisava de diseño e ingeniería y había coordinado el Data Transparency Lab, un consorcio internacional de investigación en el ámbito de la transparencia de datos. Es el fundador de MyData Hub, que forma parte de mydata.org, red de promoción de la soberanía sobre los datos personales. Es investigador asociado del Center for Organizational Innovation del Departamento de Sociología de la neoyorquina Universidad de Columbia.

Ariel Guersenzvaig
Elisava School of Design and Engineering

El Dr. Guersenzvaig es director del máster en Diseño y Comunicación, un máster oficial que da acceso a los estudios de doctorado. También dirige un curso profesional, el máster en Diseño y Dirección de Proyectos para Internet. Es profesor de grado y posgrado en Elisava, y profesor invitado de varias universidades (UPF, UPC). Tiene un doctorado en Teoría del Diseño por la Universidad de Southampton (Reino Unido), un máster en Ética por la Universidad de Birmingham (Reino Unido), un diploma de estudios avanzados en Investigación en Diseño por la Universidad de Barcelona (España) y un grado de Edición por la Universidad de Ciencias Aplicadas de Ámsterdam (Países Bajos).

REFERENCIAS

Ver listado completo de referencias en la página 15.