Scandinavian Journal of Information Systems

Volume 34 | Issue 2 Article 5

12-31-2022

Strengthening Human Autonomy. In the era of autonomous technology

Rebekka Soma
University of Oslo, rebsaurus@ifi.uio.no

Tone Bratteteig
University of Oslo, tone@ifi.uio.no

Diana Saplacan
University of Oslo, dianasa@ifi.uio.no

Robyn Schimmer
Umeå University, robyn.schimmer@umu.se

Erik Campano
Umeå University, erik.campano@umu.se

See next page for additional authors

© Scandinavian Journal of Information Systems, 2022 34(2), 163-198

Strengthening Human Autonomy

In the era of autonomous technology

Rebekka Soma
University of Oslo, Department of Informatics
rebsaurus@ifi.uio.no

Tone Bratteteig
University of Oslo, Department of Informatics
tone@ifi.uio.no

Diana Saplacan
University of Oslo, Department of Informatics
dianasa@ifi.uio.no

Robyn Schimmer
Umeå University, Department of Psychology
robyn.schimmer@umu.se

Erik Campano
Umeå University, Department of Informatics
erik.campano@umu.se

Guri Verne
University of Oslo, Department of Informatics
guribv@ifi.uio.no

Abstract. 'Autonomous technologies' refers to systems that make decisions without explicit human control or interaction. This conceptual paper explores the notion of autonomy by first exploring human autonomy, and then using this understanding to analyze how autonomous technology could or should be modelled. First, we discuss what human autonomy means. We conclude that it is the overall space for action—rather than

Accepting editor: Polyxeni Vassilakopoulou

the degree of control—and the actual choices, or number of choices, that constitutes human autonomy. Based on this, our second discussion leads us to suggest the term datanomous to denote technology that builds on, and is restricted by, its own data when operating autonomously. Our conceptual exploration brings forth a more precise definition of human autonomy and datanomous systems. Finally, we conclude this exploration by suggesting that human autonomy can be strengthened by datanomous technologies, but only if they support the human space for action. It is the purpose of human activity that determines if technology strengthens or weakens human autonomy.

Key words: human autonomy, machine autonomy, datanomous technology, robots, smart insulin pump, artificial intelligence, philosophy of technology, situated abilities.

1 Introduction

Artificial intelligence (AI) technologies "combined with automation have vast implications for organizing work and for supporting everyday life activities, but can also challenge human control" (Vassilakopoulou et al., n.d.). AI technologies make decisions and perform operations based on data-driven inferences, learning, or even self-improving (Dignum, 2018; McCarthy et al., 2006; Russell & Norvig, 1995) by gathering more data and modifying their algorithms. Today, AI technology is an integrated part of many digital systems and products, increasing their decision-making capacities (Xu, 2021). AI systems can make decisions based on data without any explicit interaction with a human.

In their synthesis of "Good AI principles" Floridi et al. (2018) present five basic principles, one of which is human autonomy. Similarly, Formosa (2021) argues that autonomy is a basic ethical principle. Floridi et al. (2018) refer to the *Montreal Declaration for a Responsible Development of AI* (2018), which states that "development of AI should promote the autonomy of all human beings and control … the autonomy of computer systems". Floridi et al. (2018) continues: "The EGE [*European Group on Ethics in Science and New Technologies (2018)*] argues that autonomous systems 'must not impair [the] freedom of human beings to set their own standards and norms and be able to live according to them'" (p. 698). In this paper we explore how these ethical principles are present in real life, i.e., we assess how autonomous technologies support or reduce human autonomy.

Autonomous systems can be defined as "systems which are capable of acting upon the world independently of real-time human control" (Soltanzadeh, 2021, p. 1; Xu, 2021). Autonomous technology can act as "tools for enhancing 'human agency, without removing human responsibility'" (Vassilakopoulou et al., n.d.). In other words, autonomous systems can have an impact on human autonomy. We are asking, in particular, whether increased system autonomy leads to a decrease in human autonomy, in a zero-sum case, or if increased system autonomy also can increase and strengthen human autonomy (Formosa, 2021; Mackenzie & Stoljar, 2000).

Questions about the relation between autonomous technology and human autonomy can be addressed in different ways. In this paper, we offer a conceptual discussion of the notion of autonomy. Strengthening human autonomy has been a long-lasting goal for IT development in Scandinavia (e.g., Langefors, 1980; Nygaard, 1996; Simonsen & Robertson, 2013) and elsewhere (Briefs et al., 1983; Soltanzadeh, 2021; Weizenbaum, 1976; Winograd & Flores, 1986). "Respecting others' autonomy is an important moral principle ... and is regarded as an important value in the design of autonomous systems" (Soltanzadeh, 2021: 16).

However, when the notion of autonomy is used to refer to technology, the meaning of the word changes (Nyholm, 2018; Purves et al., 2015).

> The suitability of the term 'autonomous' to refer to technical systems can of course be questioned. This is because current machines are not autonomous in the sense in which the term is used in moral philosophy; i.e., they do not have the capacity to reflect on and understand reasons for actions and cannot be held morally or legally responsible for their impacts on their surroundings (Soltanzadeh, 2021, p. 1).

As the number of autonomous systems increases in all areas of society, it seems appropriate to discuss what we mean by the term and how these technologies affect human autonomy. We argue that the terminology currently used to denote autonomous technologies should be more nuanced and precise, in order to create realistic expectations for what the technology can and cannot do.

The paper is structured as follows. Sections 2 and 3 discuss the concept of autonomy; section 2 looks into *human autonomy* while section 3 explores *autonomous technology*. In section 4 we discuss the differences between automation and autonomy. We rethink machine autonomy and suggest the notion of *datanomous technology* as a more precise term. In section 5 we discuss if and how human autonomy can be strengthened by datanomous technology. Section 6 concludes the paper.

2 Human autonomy

Understood in a broad sense, autonomy is "an equivalent of liberty" (Dworkin, 1988). The concept encompasses meanings such as "self-rule or sovereignty", but autonomy can also be considered as synonymous with "freedom of the will." Originally autonomy was conceptualised as an individual's capability to abide by self-imposed moral laws (Schmidt & Kraemer, 2006); the word itself stems from the Greek words for 'self' (*auto*) and 'law' (*nomos*), and can be directly translated to 'self-law'. Historically, the idea of autonomy has been tightly linked to personhood and morality. For instance, autonomy was central to Kant's conception of morality in that people are self-governing *because* they are autonomous. The understanding of morality as self-governance meant that each individual may rightly claim direction over her or his own actions "without interference from the state, the church, the neighbours, or those claiming to be better or wiser than we are." (Schneewind 1998, p. 5). In other words, individuals themselves legislate the moral law.

While morality previously was conceptualised as obedience, it changed by the end of the 18th century to be seen as self-governance, a view centred around the belief that "all normal individuals are equally able to live together in a morality of self-governance" (Schneewind, 1998, p. 4). The normative belief in the dignity and worth of the individual is the root of the more modern conception of morality as self-governance. Common characteristics of autonomy include dignity, integrity, individuality, independence, responsibility, and self-knowledge (Dworkin, 1988). The experience of respect, dignity, and equality are some of the reasons for autonomy to be valued, for example in the feeling of independence for older people still being able to live on their own in their own homes, or for young people to manage the duties of an adult citizen, paying taxes and rent. Autonomy is personal but not necessarily limited to the person, as we discuss below.

2.1 Individual autonomy is relational and situated

The individualistic view of autonomy, describing a rational man who makes his own decisions influenced by no one, has been criticized for taking an individual's circumstances and contexts for granted (Dworkin, 1988; Mackenzie & Stoljar, 2000; Mol, 2008)—particularly the support from others. MacKenzie and Stoljar (2000) introduce *relational autonomy* to denote a range of perspectives sharing the view that humans are socially embedded. A relational understanding of autonomy includes the social context that makes individual autonomy possible. In this relational sense, autonomy is an emerging property of the situation and circumstances of an individual, where his or her autonomy is encouraged or hampered by these circumstances.

People do not live in isolation but are always socially embedded, so individual humans will always be reliant on the world around them (Schmidt & Kraemer, 2006). Schmidt and Kraemer (2006) point to the paradox that autonomy requires social interdependence. A person's autonomy is better understood in the larger context of human social relations and interdependencies. Code (2004, p. 196) argues that an individualistic understanding of autonomy "glosses over the extent to which autonomous man himself is dependent on patterns of invisible advocacy" and suggests that advocacy relations are a prerequisite for autonomy (Pearsall, 1999). The notion of advocacy relations can be used to characterize a person's autonomy as one of two (Verne, 2014, 2020). The first type is an environment that supports Do-It-Yourself (DIY) autonomy by helping the individual carry out an action. The second type supports a kind of autonomy Verne calls "duke autonomy", wherein the individuals cannot or will not do the action themselves, but get somebody else to do it for them. Powerful individuals have always received support, services, and help from others in ways that do not diminish their powers or autonomy.

People might experience different and varying degrees of autonomy in different situations and contexts. The *situatedness* of autonomy refers to how a person is always in a situation. One of Heideggers neologisms, *Befindlichkeit* was formed from German *sich befinden* (meaning "existing or finding oneself in a situation") (Ciborra & Willcocks, 2006). The situatedness of any being emerges from our living in situations or contexts (Gendlin, 1978). The notion refers to how a person finds himself living in a situation, and how that situation always belongs to someone (Saplacan, 2020). Understanding human autonomy as situated enables us to see the individual human as an integrated part of his or her environment and, vice versa, the environment as a support for human autonomy. It is the situated, contextual human who acts autonomously with the support present in the situation, as well as his or her abilities. The notion of 'situated abilities' has been suggested as a way of emphasizing how both the context and abilities of a person affect his or her autonomy. Hence, the autonomy may vary over time and from one situation to the next (Saplacan, 2020). The experience of autonomy may be strengthened or weakened depending on the situation.

A relational understanding of human autonomy also includes the technologies used in the situation, ranging from common tools like glasses and canes to complex, advanced ones. Technologies that act without human command can also be understood as constituting part of a human's environment, encouraging or hampering individual autonomy. When using autonomous technology, humans may negotiate and re-negotiate their autonomy depending on the autonomy that the technology is designed to have. An example is "distributed cognition" (Hutchins, 2006; Säljö, 2001), which describes how humans design their environment to increase their capabilities by distributing tasks to, for example, a calculator (delegated the task of calculating) or a mobile phone (delegated to remember things or keep the time). Some phenomenological approaches also emphasize how we see our environment and arrange it to enhance our capabilities, e.g., how an elderly woman with mobility problems arranges objects in her home to make everyday activities easier (Brereton, 2013).

2.2 Autonomy is about space for action

Another central aspect of autonomy is the ability to decide for oneself and make one's own choices (Dworkin, 1988). Many decisions are made with little or no deliberation, like habits or automatic actions (e.g., Kahneman, 2011; Schütz, 1951). Other decisions are made after reflection, concluding a process of deliberation. Making decisions in all areas of life is part of being human, and for "anyone who lives in the real world, making choices is very much part of what makes them who they are" (Harper et al., 2017, p. 234). Making a choice implies deciding upon one of two or more possibilities (or abstaining from choosing any of them), i.e., a decision limited by a set of alternatives. When faced with an overwhelming number of choices, humans may experience decision paralysis (Anderson, 2003) and delayed or impaired decision-making, even if the outcome is of little consequence. Yet, the number of possibilities will always—to some degree—be limited by the environment, that is to say, a situation in which both humans and technologies contribute to opening up possibilities, as well as delimiting the range of options. Cohen et al. (1972) suggest that organizational decision-making depends on the participants and the problems and solutions they bring to the situation. The meeting between these elements delimits the range of available possibilities that can be chosen and acted upon, i.e., what can be decided.

Decision-making and a feeling of autonomy go hand-in-hand, because autonomy entails a certain authority over one's own decision-making (Dworkin, 1988). Autonomous humans decide what to do based on their own preferences in a situation, and this does not take place in a vacuum (Floridi et al., 2018). Basically, the autonomous human needs to have the ability to utilize the support in a situation, e.g., be able to make use of available technology. Here we focus on the decision-making and on making the right choice rather than the user's ability to utilize the technology.

Floridi et al. (2018, p. 698) argue that a "human should always retain the power to *decide which decisions to take*". Formosa (2021) suggests that autonomy presupposes "adequacy of options," that our autonomy increases when "we are given more control over important aspects of our lives and access to a diverse range of meaningful choices" (Formosa, 2021, p. 598). Hence, making a choice depends on the ability concerning the possibilities for action that are open or closed, supported or hampered, in a situation. This ability includes knowledge about the resources provided by the environment that limit and open up these possibilities. More importantly, this ability also includes the power and knowledge to influence the number and range of possibilities, and to suggest better alternatives than those originally presented. Every situation is in some ways different from all other situations and in this sense not fully predictable (Suchman, 1987). New possibilities may emerge in a situation where the resources can support new types of actions or new ways of interpreting the situation and the choices it offers. The ability needed for making decisions and acting in new situations, therefore, builds both on experience from normal situations and routines (Bainbridge, 1983; *SAE*, n.d.; Xu, 2021) and on the ability to imagine new situations or interpret the situation and its elements in new ways (Mills, 1967; Schütz, 1951).

Formosa (2021) discusses three ways in which human autonomy can be increased: 1) by being able to choose "more valuable ends" or gaining "access to a greater number of valuable ends or to ends that are more valuable" (p. 603) suggesting that it is the space of possibilities that is important for human autonomy, 2) by offering "improved autonomy competencies" through being able to focus on "build, maintain, and develop their autonomy competencies" (p. 604) emphasizing that developing new possibilities in a situation is essential, and 3) by enabling "more authentic choices, both in the sense of *more choices* that are authentic and choices that are *more authentic*." (p. 604). It is not the number of choices itself, but the range of possibilities offered that constitutes the space for action of a human.

We suggest the term *space for action* to refer to the full range of possibilities for action available to a human in a situation. The space for action is situated and can never be fully predictable. In addition to the characteristics of the situation and context, it builds on a person's abilities and imagination. An example from fieldwork in a hospital can illustrate the difference between an expert and a novice (Bratteteig, 2004; Goodwin, 1994). The authors followed the nurses as they were doing their work. After a morning round with the doctors, the nurses asked whether the researchers noticed how they lifted the patients' ankles and patted them on the leg. The nurses explained that looking at and feeling the ankles is a way to check if the heart functions well. What the researchers had interpreted as care for their patients was in fact a professional evaluation of the patients' condition.

The human space for action is, in other words, characterized by the ability to redefine a situation. When changing or adding perspectives, a person may understand the currently available information in new ways. In turn, this introduces a new set of possibilities for action, alternative to those originally available. The quantitative change in the choices available is trivial, as it is the qualitative nature of the new set of actions that redefines the situation. A good example of this is the story about Odysseus, who voluntarily agreed to give up his autonomy and be tied to the mast, so he would be able to withstand the sirens' song. In the story, Odysseus redefines the situation from *it is not possible to listen to the sirens' song without suffering the consequences of its enchantment, which is certain death* to *it is possible to listen to the siren song, as long as I am prepared to withstand the suffering of being tied to the mast.* In allowing himself to be tied up, he temporarily loses all autonomy but does so precisely in the pursuit of gaining autonomy. This illustrates how the space for action has a variable horizon that can be extended or constrained, depending on what the situation calls for.

A relational and situational perspective on human autonomy emphasizes that autonomy as a concept cannot be understood as a dichotomy. We cannot say that an individual is either autonomous or not; her or his autonomy will vary with the characteristics of the situation (Mackenzie & Stoljar, 2000). Self-driving cars can illustrate this point. For a person with impaired vision a self-driving car will be of assistance in running errands or traveling places, without relying on the help and kindness of family and friends. This can give the visually-impaired person more autonomy. On the other hand, an experienced driver, who enjoys operating an ordinary (non-self driving) car may experience her autonomy as limited or constrained in a self-driving car. The pleasure of being the one in control of the vehicle—such as pushing the limits of the car by speeding and changing gears—is taken away from the driver with this technology, and she would be more like a passenger under the control of the car. Therefore, the same type of technology may be experienced in different ways. A self-driving vehicle has the potential to both weaken and strengthen human autonomy depending on the vastly different circumstances of the potential driver (or passenger).

3 Autonomous technology

It is generally accepted that, in delegating a work task to autonomous technology, humans are relieved from, for instance, tedious, repetitive, or dangerous work. Technology is commonly considered as autonomous when it appears to make decisions independently of human intervention, that is to say, without explicit human control or interaction. In more precise terms, a machine is autonomous when it is capable of following "a complex algorithm in response to environmental inputs, independently of real-time human input." (Formosa, 2021, p. 599). In other words, autonomous technology makes decisions on its own and its decisions can later be changed based on the gathering and processing of new data.

A central issue for the development of responsible AI (and autonomous technologies) is the question of what it actually means for an AI system to make a decision, and whether or not such a system can be held accountable for its decisions and actions (Dignum, 2018). Soltanzadeh (2021, p. 1) contends that autonomous systems "cannot be held morally or legally responsible for their impacts on their surroundings."

Furthermore, Bradshaw et al. (2013) criticise the common view of autonomous technology differentiating between levels of machine autonomy (Cummings, 2004; Parasuraman et al., 2000; *SAE*, n.d.) and argue that the concept *autonomy* refers to different phenomena. They particularly discuss self-sufficiency and self-directedness, i.e., both "the capability of an entity to take care of itself" and its "freedom from outside control" (Bradshaw et al., 2013, p. 54). Both self-directedness and self-sufficiency are needed for a technology to act autonomously.

3.1 Seemingly autonomous behavior

"[E]ven the simplest machine can seem to function 'autonomously' if the task and context are sufficiently constrained" (Bradshaw et al. 2013, p. 57). For instance, the thermostat exercises "an admirable degree of self-sufficiency and self-directedness with respect to the limited tasks it's designed to perform through the use of a simple form of automation" (Bradshaw et al 2013, p. 57). A thermostat regulates the temperature automatically. A smart building that regulates the temperature autonomously to save energy is based on a complicated balancing of various types of data—possibly including that of a thermostat. The building's computer may decide not to adjust a low indoor temperature should other data, e.g., date or time of day, be given higher priority in the calculation.

Knowing both how the smart building's sensors are set up and also its energy-saving model makes it possible to understand how it operates. However, this is not the case with all autonomous technologies, even everyday and mundane technologies like robot vacuum cleaners. A machine's self-sufficiency will in most cases be relative both to a set of pre-defined goals and tasks as well as a set of fixed ones. This means that even when a system functions well and operates immaculately without human supervision, it all happens within the limited circumstances of its abilities (Bradshaw et al. 2013).

Take robotic vacuum cleaners as an illustration. When turned on, a vacuum cleaner robot moves forward and collects dust and small things from the floor on its way[1]. Sensors detect obstacles it has to climb. The robot is unable to differentiate between a chair, a foot, or a ledge, and thus cannot have a variety of strategies for how to deal with such an array of challenges. Whenever a obstacle is detected, the robot backs up a few inches, turns itself some degrees to the side before continuing forward, collecting dust until the next hindrance is met. In this way, the robot covers the ground through a seemingly unpredictable series of movements. When its power gets low, the robot returns to its charging station, which is equipped with sensors that the robot can detect. Some vacuum cleaner robots use an infrared sensor to create a map of the navigated environment. Such maps are made available for the user on a corresponding smartphone app. Other than the app, a user has little control of the robot during its operation, other than turning it on or off. Its movements may also be manipulated by positioning an obstacle in front of it (e.g., a foot) to force a change in direction.

Empirical studies of robot vacuum cleaners show that the robot takes over some work tasks, but also requires extra work from human users. Even a simple task like vacuum cleaning, delegated to a robot does not make the work disappear; it just changes the overall activity (Bradshaw et al., 2013). Robots in the wild (i.e., natural or human environments not specifically designed for robots) require humans to arrange the environment so that it fits the robot's requirements for performing its tasks autonomously (Forlizzi & DiSalvo, 2006; Verne, 2020). For example, many users report that they continuously move furniture around in their homes. The various kinds of facilitation work required to enable a domestic robot to carry out its task autonomously can be described as pre, peri, and post-facilitation work (Soma et al., 2018), and comes in addition to the initial setup of the robot. Pre-facilitation includes removing things lying on the floor or moving furniture around. Peri-facilitation is carried out while the robot is operating, e.g., helping it when it becomes stuck in cables, curtains, or items users forgot to remove before the robot started. Post-facilitation takes place whenever people make permanent changes to their homes to relieve themselves from pre and peri-facilitation (Soma et al., 2018). Of course, we normally also do facilitation work before cleaning, whether we do it ourselves or hire a cleaning service. Although the basic task of tidying before human or robotic cleaning assistance may look identical from the outside, they are very different in terms of intentionality and outcome. In the unfortunate event that you have forgotten to tidy a room or forgotten to pick items up from the floor, the housekeeper is able to do this work for you or downright refuse work under such conditions. They will not, however, use the vacuum cleaner and tangle it in the things you missed or forgot to tidy. In contrast, the chances of a vacuum cleaning robot successfully performing its task without proper facilitation work is slim, because it is unable to differentiate between obstacles.

Even if users carry out facilitation work before, during, and after the robot operates, many still conceive the robot as autonomous. Our studies indicate that the fact that it moves around based on its internal decision-making—avoiding obstacles and operating without direct human command—is what makes many people ascribe autonomy even to this type of simple robot (Soma, 2020). Merely collecting dust is automated by the robot. Hence, the robot vacuum cleaner is a simple automaton, and in this sense comparable to a washing machine. However, it is not trivial that the internal decision-making process of these robots—where to go, when to stop and turn or not turn—is opaque to the user. Autonomous movements through physical and social space are easily interpreted as the robot having the capacity to act, react, and even adapt to its environment (Bianchini et al., 2016). It is the way it goes about doing it, moving by itself, that makes us attribute reason and agency to the robot. In other words, it is disguised as an autonomous technology, moving around on its own, following a self-devised plan for how to best ensure clean floors, without human command.

This suggests that the vacuum cleaner robot can only be seen as autonomous if we focus on the technical artefact in isolation, viewing autonomy as a property or feature of the technical artefact itself, ignoring the facilitation work that is performed around it (Soltanzadeh, 2021).

3.2 Autonomy or automation?

The autonomy of technology is difficult to identify as it is often confused with automation (Xu, 2021). The concepts 'automation' and 'autonomous technology' are often used interchangeably, but they refer to different technical capabilities. Automation means that the performance of a task is delegated to a machine, like washing clothes in a washing machine. Simple automatic machines such as washing machines or ATMs present a set of choices to the user. However, there are also automatic machines that operate without presenting information or choices to the user after deployment, during operation, e.g., a pacemaker. Human decision-making and control depend on being presented the information that the decision is based on and being able to make the decision oneself. Autonomous technology does not offer this possibility, and its control has to be achieved in different ways, such as when you stop the robot vacuum cleaner by putting a foot in its way so that it turns around. When it comes to autonomous systems, the initial technical premise upon which the user decided to take the system into use changes over time, during its operation.

Above, we discussed whether the robot vacuum cleaner is in fact better understood by the notion of automation, because the only decisions it makes are tied to how it physically covers the surface area of a given space and turn when it detects an obstacle. We suggested that its movements seem meaningful to a human observer, making the technology appear autonomous. There is not just an absence of human command; there is also a certain lack of interaction.

Lack of conscious interaction from the human side in operating automatic technology is quite common, e.g., automatic door opener. However, such "implicit interaction" (Ju & Leifer, 2008) does not distinguish between an autonomous and an automatic system, in which the system's data gathering detects human action that can be interpreted as a command, e.g., approaching an automatic door and get registered by the sensor, which is interpreted as a command to open the door. Ju and Leifer (2008) give an example of a high level of implicit interaction. A friend sends a link to a funny animation online, but your computer needs to install a plug-in for you to view it. According to the authors, this plug-in can be installed in several ways. The highest level of implicit interaction is described in this way:

> Our Web browser anticipated that we might want to play a Flash animation someday, and already has downloaded and installed the plug-in. … This … case is the most implicit interaction. In fact, with so much presumption and so little visibility, this … interaction may hardly be considered an interaction at all, since there is no activity or awareness on our part. … [I]t is more accurate to speak of interactions being more and less implicit. (Ju & Leifer, 2008, p. 77).

Ju and Leifer characterize this interaction as proactive, which means the technology takes the initiative. In addition, it operates in the background, requiring no attention from the human user. This example allows us to distinguish between automatic and autonomous computer actions. The former merely install an update of software on a user's computer, while the latter initializes the search for new software *by itself*—new software that the user might use in the future but has not yet used, and never asked for. Machine Learning (ML) adds to the complexity by integrating new data into its basis for decisions, and the original rationale is replaced by machine logic after some time (Holmquist, 2017) making the decisions less transparent and, in turn, more difficult to control. Autonomous technology makes decisions based on data it has collected. These data—and decisions previously made upon them—is used to modify its model of the world.

3.3 AI in autonomous systems

In the start of this section we quoted Formosa (2021)'s definition of machine autonomy. Formosa continues by saying that "[m]ore advanced forms of machine autonomy typically depend upon the use of AI" (Formosa, 2021, p. 599). AI makes decisions and performs operations based on data-driven inferences, such as rules (e.g., chess rules applied to decide on the next best move) or statistical calculations (e.g., a spell checker that suggests the word you most probably are trying to type) (Holmquist, 2017). ML learns from the new data it gets exposed to and improves its decisions over time.

An example of a simple AI-based autonomous technology is a smart insulin pump which constantly and automatically measures glucose levels and administers insulin in micro-doses throughout the day. In a study of how a smart pump is used in real life we learn about Lisa, who recently shifted to a smart insulin pump[2]. The smart pump replaces—and eliminates—previously needed equipment such as the blood glucose meters used to determine the proper insulin dose and the syringe or pen used to inject the insulin manually. The smart pump includes a constant glucose monitor (CGM) sensor that feeds a computer with measurement data, which in turn automatically controls the dosage and distribution of microdoses of base insulin. In other words, it is 'smart' in the same sense as a 'smart' building. It combines specific data concerning blood sugar levels and calculates the right insulin dosage.

Injecting a wrong dose of insulin will cause the patient to become unwell, and both underdosing or overdosing the amount of insulin injected may be life-threatening. When well trained, a smart insulin pump is less error-prone than insulin delivery by humans. Thus, the smart pump is an example of a medical technology that makes a complicated but life-saving treatment easier to deal with on a daily basis. However, due to the highly individual differences between each diabetic patient, a new smart pump does not work right out of the box. Instead it has to learn from the daily life and routines of its user in order to adapt its algorithms for automatically distribute the right dosages in the right situations. The setup procedure for Lisa's new device and its learning process made it necessary for her to clear out her work schedule, travel plans, and other appointments, while calibrating the device to her body specifically. Should she radically change her life style, she would need to recalibrate the pump.

In Lisa's case, the autonomous pump has reduced her control of the insulin administration compared to her previous manual pump. Blood glucose levels are measured continuously, but the insulin dose response will always be after-the-fact, after a rise in the glucose level. In this regard, a smart insulin pump is better at imitating the self-regulating hormonal feedback-loop in a healthy and fully functioning pancreas, similarly to how a pacemaker imitates the beat of a healthy heart.

The level of control constitutes one of the main differences between the two generations of pumps. Lisa cannot herself add contextual factors that would alter the insulin supply. She does not fully understand how the pump works, and neither does her diabetes nurse. They have to trust the pump's calculations, and the trust is based on the experience of living with the pump. To maintain trust in the pump Lisa has to 'behave correctly'. Her control of the pump lies in her adherence to the model of glucose-level control that the pump uses to process its data and build its decisions on.

The smart insulin pump is an example of a technology that is equipped with a ML algorithm specifically developed for a limited purpose, namely to make it easier for diabetes patients to deal with their condition, hopefully enhancing their overall quality of life. The pump constantly learns from the data it collects and modifies the insulin doses to be better tailored to the individual, so its learning is limited to the individual to which it is attached, and its autonomous decisions are thus contained within a very limited system.

4 Rethinking machine autonomy: Datanomous machines

As a result of the constraints on our thinking with inadequate language, many misconceptions arise as to what an autonomous technology is, can, or should be. As we have explained above, the concept of autonomy has deep roots in moral thought. Keeping these roots in mind, autonomy implies that an entity described as autonomous has a right to, or possesses a capability for, self-governing. Still, it seems that the concept's deep roots in moral thought are no longer prevailing, as autonomy usually refers to independence in modern dictionaries. Thus, it has gained a far more functional meaning in the vernacular use of the word. In everyday language, autonomy is used as an adjective, as a characterization or dimension of being, rather than a state of being that contains a multitude of dimensions.

According to Luck et al. (2003) an autonomous agent is an entity that is able to generate its own goals. Still, the authors accentuate that autonomy as a concept is "one of the most used but least operationalized words in the fields of intelligent agents and multi-agent systems" (Luck et al., 2003, p. 9). It is common to use autonomy as a characterization of a type of technology, usually expressed through neat levels that easily sort the autonomous from the regular systems. Such views have been criticized for their descriptions of autonomy and automation as being one-dimensional, a misconception which is not just damaging in its own right, but can, through its continued propagation, lead to serious consequences (Bradshaw et al., 2013).

The current use of autonomy in these fields might therefore be considered a misnomer. The currently functional meaning of the term, combined with the lack of operationalization, leads to a confusion of what the notion contains. Therefore, Schmidt and Kraemer (2006) suggest that it might be fruitful to reintroduce the original, humanistic meaning of autonomy.

In a Heideggerian transcendentalist view, language is understood as a condition for the possibility of thinking (Coeckelbergh, 2017). While neither language nor technology determine us, per se, the language used to conceptualize technology—both novel, well-established, and fully integrated—might provide unintended constraints on our thinking, constraints of which we are usually unaware. The roles of language and technology as mediators mean they shape our way of thinking. Thus they also shape our actions by framing our understanding of the world in a certain way. Heidegger's approach to overcoming the inadequacy of language to properly describe the phenomena he was interested in examining, was to create new words. This enabled him to understand the world and technology in new ways. Inspired by Heidegger's strategy we wish to introduce a new term that reframes the way autonomous technology is conceptualized. Our hope is that this new term will reshape how autonomous technology is understood, not just within research but also for users of this technology.

As mentioned above, autonomous systems depend on data to function properly. Data is perhaps best understood as an indicator or representation of a phenomenon, like fever indicates illness. In modern dictionaries, data usually denotes representations of items of information in the form of sets of values or variables, collected together for reference, analysis, or calculation. This definition rightfully indicates that data is purposefully curated, and intended to assist in identifying the meaning represented by the values. In other words, data is designed (Feinberg et al., 2017; Muller et al., 2019). Studies concerning the work practice of data scientists show that "data is produced by techniques of measurement that are imbued with judgements and values that dictate what is counted and what is not" (Pine & Liboiron, 2015). Further, data come into existence as given; as captured; as curated; as designed, and as created (Muller et al., 2019, p. 1) with increasing levels of intervention by data scientists. This practice of data wrangling filters out dirty data that do not fit, which in turn indicate that the system that uses this data influences "the kind of data that [can] be represented" (Muller et al. 2019, p. 3). In other words: "Raw data is an oxymoron" (Gitelman, 2013).

When new data is collected, the original model of an autonomous system will become more precise in representing the world according to its specific data types. Collecting and integrating more data in the basis for decisions is the system's way of expanding its space for action. For instance, Lisa's smart pump becomes more precise in giving more adjusted dosages of insulin the more it knows about Lisa and her glucose and insulin balance, because this can be measured and represented as data. An automaton operates on a model with data that does not change its range of actions. Autonomous systems can change their range of actions when more and new data is integrated and when decisions change due to new data processing. Lisa's smart pump is an example of this, using data to analyze patterns of glucose and insulin levels over time. These data make up the basis for adjusting the insulin dosage at the right time. Similarly, a smart building regulates the temperature, and electricity, based on various measurements as well as information about time of day, weekday, season, and so on.

The insulin pump is a computer system that makes inferences based on the data it is able to collect and the calculations it is able to do. The system does not know anything about the phenomena it measures or the operations it performs. Just like a robot vacuum cleaner does not know anything about dust, the insulin pump does not know anything about blood or insulin. As humans we can make sense of the data, e.g., that a piece of cake will change the glucose level, or judge whether a thing on the floor is dust or something we want to keep. The systems' performance, or their space for action, is limited by the data they can collect because data constitute the basis for the decisions the systems make.

As all smart insulin pumps are tailored to individual patients, they are all different. Further, it is smart only within a closed system, similar to a thermostat. The pump expands its possibilities within the closed system and can develop many more possibilities than a human in that space—similar to a chess computer that reasons differently than a human within the limits of the game (Holmquist, 2017).

The robot vacuum cleaner acts on its data by turning when it senses an obstacle; it does not matter what the qualities of the obstacle are, e.g., whether it is big or small, soft or hard, blue or yellow. An obstacle is an obstacle. The smart pump can change the insulin dosage based on its input and its calculations, but its actions are still limited by the data it can register and what it can process. Literature on ML often refers to playing Go and chess. The computer has implemented the rules only, and finds ways to play by itself. It creates its own decision logic (Holmquist, 2017). This is possible because both Go and chess are closed areas, different from many other real world activities—like vacuum cleaning and diabetes management, in which the data limits the context. We have argued that both the robot vacuum cleaner and the smart insulin pump operate in ways that show their limits compared to human perception of data. While the robot vacuum cleaner operates with a stable and very limited set of data in its world, the smart pump learns from new data in ways that modify its set of data and hence its decisions. It behaves autonomously.

However, both systems operate on pre-defined types of data based on human categories. The data types are limited; the types close the system. Even an AI-based technology that learns from new data can only act on the data types on which it has been trained. Understanding the data—and the types of data—is therefore crucial for understanding these technologies. Data is what the sensors or other data gathering equipment are able to register, i.e., data is a result of processed input, and a product of design. This production of data has to happen before it is possible to compile and process the data. Thus, data is always historical. Only pre-defined types of data will be collected or taken as input, which makes it difficult for data-driven technologies to adjust to new situations (Broussard, 2018). Therefore, the particularities of a situation will not be accounted for, unless they are part of the pre-defined data set.

Computers use data as input for their algorithms to operate. Hence, it is the data that governs the technology's decisions and behavior. It is therefore difficult—even impossible—for a computer to increase its space for action beyond the pre-defined data categories and algorithms unless it operates in a closed area like chess or Go (Holmquist, 2017), whereby the ML technology can alter its algorithms within the limits of the game. Any space for action created by data types and algorithms can be seen as a model. It is not changed (expanded or constrained) even when more data is added and used to make slightly different decisions. The decisions are still limited to the modeled space, because data is still of the same specific *type*. Adding more data will contribute to increasing the detail and granularity of the modelled space, but it does not contribute to its redefinition.

In contrast to the autonomous human, autonomous technologies do not have situational awareness. When they are developed, they are equipped with a model of the world that is compatible with the desired or available data types. Any and all operations of these systems will be limited to the world-view offered by its specific data set. In turn, all its decisions will be driven by registered data, i.e., the data is historical. This data is either the input that it has been fed by human trainers, or retrieved itself in its interactions with its environment. While such systems will operate in many situations without human intervention, these technologies are unable to adapt and operate independently should the situation change.

It is obvious that there is an inherent difference between what we consider to be human autonomy and the limited operations of autonomous systems. We argue that there is a need to make these differences explicit, and bring forward a new and more precise concept to assist this process: *datanomous technology*.

As a term, datanomy (and datanomous) suggests that the operations of these technologies are responses to their available *data*, and not the *situation* at hand. Viewing the technology as datanomous, rather than autonomous, makes explicit that it is not governed by a self, but by data. To exemplify, we will substitute the self with data in some of the dimensions that characterize autonomy. *Self-sufficient* and *self-directed* becomes *data-sufficient* and *data-directed*.

As we have noted above, the data on which the vacuum cleaner robot operates as it moves around the home is very limited. The robot has no understanding of dust or cleaning. It registers obstacles but does not distinguish between humans and furniture, between cables and dirt, or between an apartment and the hallway of an apartment building. Obstacles that cannot be handled result in a shift in direction. The robot also does not register smaller things located on the floor, and may easily become stuck in carpet frays or tangle itself in cables. Parallels can be drawn to a similar technology, the robot lawn mower, a technology that also does not have any situational understanding of what it is doing. It does not know about grass or gardens, and will not change its path for small objects on the lawn before it is too late and the objects obstruct its operations (Soma & Herstad, 2018; Verne, 2020).

Even in a less-than-ideal environment we may say that the users assisted the robot in becoming data-sufficient by facilitating it in its operations. Similarly, a vacuum cleaner robot is designed to navigate in a home environment, its cluttered nature means the robot will always meet unpredictable situations, different from those it was created to handle. Its data-sufficiency and data-directedness will *always* be limited by its data. The robot is neither data-sufficient to carry out the navigation task when the environment is unpredictable nor when it is not programmed to handle the myriad of furniture, people, animals and other objects that can act as obstacles in the home.

Similarly, the smart insulin pump was designed for a very specific purpose, namely, to adjust insulin levels based on the data it receives. If we limit the smart pump to responding to glucose measurements—a closed system—we can say that the smart pump is indeed data-sufficient. The pump operates on the specific data with which it is trained to interact, and responds with micro-doses of insulin if the glucose level measured in Lisa's body is too high. In this sense, the device is quite simple, and it is easy to see that the scope of its data is limited, even if the insulin dosage calculations are complex and inscrutable. The calculations include patterns of glucose oscillations over time, calibrating the model originally implemented in the device.

However, if we widen the scope to include Lisa and her life, situations and contexts constantly change. The smart pump can never understand these. In the larger perspective of her life, the pump is not data-sufficient. For instance, if Lisa attends an event where cake is served, the pump will not know or understand this new situation; it delivers insulin based only on historical data. As a general rule, Lisa can no longer act preventively by increasing the dosage before the glucose level rises, as she could do with her old device. This would provide wrong data that would lead to wrong calculations. Everyday use is also always training (Holmquist, 2017). The smart insulin pump responds with insulin on continuous measurements of the body's glucose level. It is not a device with which you can negotiate.

ML-based systems improve over time. Their enhanced functioning and expansion of their space for action come with the increased amounts of data providing a higher resolution or granularity of data on which to base decisions. However, the datanomous technologies are still limited by their data compared to a human space for action, where fundamentally new and unpredictable types of data can be incorporated into the decision-making. Datanomous technology is based on existing classifications, making the limits of classification systems important to acknowledge (Bowker & Star, 2008; Broussard, 2018).

Datanomous technology interacts with humans through their data. There is often a lack of interaction mechanisms for the human user to control the technology. This means that the user of a datanomous technology is at the mercy of the designer's understanding of a future and very specific use situation. Even ML-based technology is framed by the original design of the algorithm and its data types, which forms the entire basis of the decisions it makes.

5 Strengthening or weakening human autonomy

In this paper we have argued that human autonomy does not depend on controlling the technology, per se. Human autonomy is instead correlated with understanding the technology's operations and requirements. People can choose to facilitate the environment to better fit the technology by improving the conditions for datanomy. In so doing, they increase the chances that the datanomous technology operates as intended. This, in turn, may strenghten human autonomy.

5.1 Understanding rather than control strengthens human autonomy

Throughout the paper, we have explained how knowledge of the datanomous technology and its data can be more important for human autonomy than direct control over the technology. We have argued that the new concept, *datanomous technology*, captures this insight. The knowledge of and ability to use a rather incomprehensible or imperceptible technology have to be based on trust that the technology functions as expected, i.e., that its operations are correct and predictable. A user interacting with datanomous technology gains experience with how it works. This experience can be used to explain its behavor, if the technology consistently manages to fulfill the tasks it has been delegated. Users may further experience that the technology makes good choices for them or assists them in making good choices for themselves. Giving away control to imperceptible technology can be compared to trusting a doctor's expertise, and that the doctor's advice is in a patient's interest. Yet, it is important to note that trust in a technology in reality means trusting the people who designed the technology, its algorithms, its data and data types, its way of collecting new data, as well as those who worked with the data (Muller et al., 2019). Similarly, when we are in need of medical treatment, we trust the doctor prescribing treatment because we trust the institutions that guarentee that this person *is* in fact a doctor with the necessary qualifications for treating you.

Being familiar with the data and algorithms in datanomous technologies gives the user a better basis for choosing how to use the technology as well as how to relate—or work—with that data and the datanomous environment (e.g., Haapoja & Lampinen, 2018). It is a paradox that having to do facilitation work before using the robot vacuum cleaner can be seen as leading to more human autonomy, as it has the potential to teach the human user about the limits and possibilities of the technology.

Technology that is imperceptible is often used in ways suggested *by* the technology, such as following certain steps and applying default values. Systems that narrow the decision space or try to *nudge* the user into a particular direction are often perceived as patronizing or detrimental to subjective autonomy (Thaler & Sunstein, 2008). Nudging can be seen as hints to support the datanomy of the technology, and we argue that this may also in fact lead to human autonomy, but in a different sense. Nudging is particularly helpful when the user does not have the ability to make good choices for himself, i.e., when systems are too complex and complicated or when the interaction is so implicit that it is imperceptible, like many of the autonomous systems integrated in public and private services. Having a higher number of choices does not automatically lead to more autonomy (in particular, if you do not understand what they mean), just as fewer choices do not necessarily lead to less autonomy. It is more complicated than this. Just like in the story about Odysseus, taking away your freedom of choice for a limited time will not necessarily mean that you become *less autonomous* overall.

We have argued that people have limited knowledge of the technology and how it works. Therefore they have to rely on their experiences from interacting with the technology over time, for them to form a trusting relationship with it. This may not be problematic when the technology is small and its scope limited, but in complex systems with several smaller interdependent and interacting parts it becomes an impossible task to gain an overview over the vast number of operations. Examples of such systems can be found in government, health, infrastructure, and media. Individuals using or being reliant on these kinds of systems cannot know whether they encounter all or just a part of the system, nor which part. Even for skilled users, it will be impossible to gain an overview over what is recognized as data by the system. Furthermore, it is impossible to know exactly how this data is used, and for what purpose. In these cases, users will not be able to know if they comply with the demands required by the system for it to strengthen their autonomy and widen their space for action.

AI complicates the matter more. In a paper discussing how machines 'think', Burrell (2016) argues that there are three forms of opacity in ML algorithms: 1) intentional, as when a company wants to protect its product, 2) the technical illiteracy of their users, and 3) the specific characteristics of ML algorithms "and the scale required to apply them usefully" (Burrell, 2016, p. 1). Burrell is particularly concerned with "the mismatch between mathematical procedures of machine learning algorithms and human styles of semantic interpretation" (Burrell, 2016, p. 3) referring to the layers of statistical analysis performed on data to calculate the probability of which category a piece of registered data belongs to, as well as the process of yielding a conclusion from that analysis. She illustrates her point with spam filters, which operate based on "a matrix of weights that will then be used by the classifier to determine the classification for new input data … for example … emails that have been pre-sorted and labeled as 'spam' or 'not spam'" (Burrell, 2016, p. 5). She shows that the filter collects a 'bag of words' used for classification of emails as spam, which is different from a semantic analysis of spam emails. As the number of words increases—as the ML learns from new data—the granularity of the filter's classification becomes finer and finer and more and more difficult to understand. "[I]ntuition fails at high-dimensions" (Domingos, 2012) of statistical analysis. "[R]easoning about, debugging, or improving the algorithm becomes more difficult with more qualities or characteristics provided as inputs, each subtly and imperceptibly shifting the resulting classification" (Burrell, 2016, p. 9). This imperceptibility into the inner workings of datanomous systems contradicts the possibility for true Explainable AI.

In making our argument we have described simple and limited autonomous systems in which the user can directly see or sense the system's operations. In these systems, it is clear both *what* its data is and *how* that data is used in the system's operations. In large systems, however, composed of many smaller subsystems which interact with each other, comprehending the system's operations becomes difficult. Concrete situations such as Lisa's case make the challenges to human autonomy from imperceptible technology accessible for scrutiny. For instance, human autonomy as a tenet within bioethics points to the right of the individual to decide whether or not receive medical treatment (Floridi et al., 2018). In Lisa's case, it is possible to evaluate whether she is truly able to consent to the treatment given by the pump. The pump's workings become increasingly opaque as its algorithms adjust themselves to the data they register about her insulin regulation needs. The fact that the pump is a small, manageable, isolated system, and that Lisa can identify the effects of her behaviour on the system and the system's developmental response (e.g., that she needs to provide the system with correct data), allows for an evaluation of her autonomy when using the datanomous technology and the treatment it provides her on a personal level. However, the complexity rises at an organisational level comprised of many different situations, all with their particular contextual relations constituting the human autonomy. It therefore becomes next to impossible to assess (in a general way) the effects of a large system on human autonomy. Yet another layer of complexity arises whenever systems 'learn' and change the data on which they operate over time, with these changes originating from data processing (Burrell, 2016) rather than a context understandable to a human user.

We have argued that when humans understand the data, it is possible for them to choose how to use the datanomous system in a manner that strengthens their individual autonomy. If the data is opaque, this possibility is reduced, and the larger system operates with the datanomous system as its core, rather than strengthening the human-in-context. Human autonomy and machine datanomy are not mutually exclusive, i.e., humans do not automatically gain or lose autonomy when datanomous technologies are implemented. Still, it is important to remember that humans and machines are radically different when it comes to their possibilities of being morally and legally responsible for their actions (Soltanzadeh, 2021). *If*, *when and how*, and *for what* the technology is used, determines whether human autonomy is weakened or strengthened.

5.2 The activity creates a context for datanomy

In discussions concerning whether an activity or a task is better delegated to a datanomous technology or not, we need to examine the *nature* (which can be either instrumental or constitutive) of the activity and the *role* (either epistemic or logical) of human engagement (Soltanzadeh, 2021). In activities where it does not matter *who* performs the activity or *how* it is completed, such as driving between work and home, the activity can successfully be delegated to a driver or a self-driving vehicle. However, people also perform activities for which the value of the activity is in the engagement itself (Carr, 2015). Examples are abundant, like participating in a play, playing a game, engaging in religious acts, or listening to music, to name a few. Whenever the desire is to engage in such activities for the sake of engagement, these specific activities cannot logically be delegated to others, neither human nor machines. Driving cannot be delegated to another driver or a self-driving car if the goal of the driver is recreational. When engaged in leisurely driving, what constitutes a good, relaxing, pleasant, or exciting drive can only be evaluated by the individual. This way of driving cannot be planned in detail and depends on a series of personal, momentary decisions, and it is therefore neither logically nor epistemically possible for a datanomous technology to be delegated this activity. What further complicates the matter, is that when merely observing someone drive, it may be impossible to decide whether the driver is constitutively engaged in leisurely driving, or merely driving to reach the instrumental goal of arriving somewhere. It may even be both.

This distinction can be used to discuss both the robot vacuum cleaner and the smart insulin pump. At first glance, removing dust—when the successful result can be evaluated by a clean floor—is obviously an instrumental task and can therefore be delegated to a robot. Hence, it would be within reason to claim that this type of activity can be delegated to a datanomous technology, which it has been. However, for some people, the act of contributing to a tidy and clean home may be—in itself—satisfying (Christiansen & Andersen, 2013). In such a case, value can be found in the experience of performing cleaning activities. Since delegating the task of removing dust can be delegated to a robot, it needs facilitation to be successful in achieving this instrumental goal, and a person could still be engaged in the cleaning activities even if the physical dust collection is not one of them.

The smart pump is more complex. The device is developed to deliver a specific and measurable result: a correct insulin dose to the body. However, it is important that this does not eclipse the fact that living with the pump is a personal experience, whereby the overall goal of coping with diabetes can be considered a value. A well-functioning smart pump will make living with diabetes—performing insulin managing activities—easier. It is still worth noting, however, that the active management type of activity is transformed into an instrumental activity when performed by the pump rather than Lisa herself. Without judging what is better or more desired, we find that the competence

Lisa had as an active manager of her diabetes—keeping track of the effects of the insulin dosages—may disappear as the management has been delegated to the datanomous and black-boxed technology. Still, Lisa has ended up very satisfied with her new device. Like Odysseus, she gained a new kind of freedom by delegating the management of diabetes to the smart pump. Even if the pump in some ways controls her life, she worries less about her health.

In light of this, it may be worthwhile to ask if there are any activities that may, over time, have unforeseen societal effects—for worse or better, for a strengthened or weakened human autonomy—of forcing a transformation of constitutive activities into instrumental ones. Consider activities that are not obviously constitutive, because they are industrious and exhausting, or dull and repetitive, yet could also be considered valuable, such as caring for children or elders. In the case of the smart pump, the knowledge of routine management of diabetes may be lost over time when Lisa no longer must—or can—control the pump. Delegating boring routines to automation may result in loss of knowledge about these routines, making it difficult or even impossible to handle non-routine situations or crises (Bainbridge, 1983; Xu, 2021). In delegating routine parts of a task to automation the person in charge of facilitation and articulation work could end up without any control over the quality of this automated task. Therefore, it only makes sense to use automation in activities that are possible to describe in steps of objectively identifiable results (Soltanzadeh, 2021). Otherwise, the process of automating the activity necessitates its redefinition. For instance, when the activity of 'looking after' a patient is delegated to monitoring sensors and cameras, the aspect of 'care' is completely severed (Bratteteig & Wagner, 2013). Automation does not contain discretion or care and should not be used whenever these qualities are an important aspect of the activities.

We ended section 3.1 by stating that *if*, *when and how*, and *for what* the technology is used, determines whether human autonomy is weakened or strengthened. This fits well with Dignum's (2018) insight about the role of values in protecting human rights, emphasizing that values depend on the socio-cultural context in which they are adopted; the individuals adopting them; and how they are prioritized. Organizations are made up of individuals; it is individuals who carry out tasks and make decisions and realize values in practice. The goal of every activity therefore becomes closely tied to ethics, legislation, and responsibility—including when technology is part of that activity (e.g., Mikalef et al., 2022). The responsible organization facilitates the individuals' possibilities for acting responsibly, also when the activity includes datanomous technology.

6 Concluding remarks

In this paper we have explored whether datanomous technologies can strenghten human autonomy. We have described human autonomy as the space for action in a situation, i.e., the capacity to act according to one own's values and wishes utilizing the resources in the situation. Human autonomy is relational and situated. We introduced the notion of *datanomous technology* to denote technology that acts in a way that is interpreted as autonomous. The concept datanomous emphasizes the role of data as the basis for the actions that the technology is capable of, and that even if new data is included in the decision-making, the data types limit its space for action. Datanomous technology is not able to add data types or redefine its model beyond an increased resolution.

Based on the exploration of the notion of autonomy above we can arrive at a better understanding of our use of the notion for humans and technology. Human autonomy has to do with people's ability to manipulate their space for action. In knowing about possible actions, understanding how to generate different possibilities by making use of the available resources, the space for action can be altered. In the case of Lisa and her smart insulin pump, we saw that even if she has given up understanding and controlling the device, she still feels that it enhances her autonomy. In the case of the robot vacuum cleaner, we saw that even if the users carried out a lot of facilitation to make it work, they still interpreted it as 'autonomous' technology that helped them with their cleaning activities. Technology—whether datanomous or not—obviously can become a part of human autonomy.

We have discussed the concept of autonomy by specifying what it means, arriving at an understanding of 'autonomy' that is best described as containing a multitude of dimensions rather than a spectrum. We have argued that 'autonomy' is *not* a characterization, feature or property, but better understood as a condition or state of being. The fact that autonomy is often used as a characterization (for both humans and machines), has contributed to its meaning becoming confused. In our understanding of the concept, the most important dimension of autonomy is an individual's space for action, which frames the ability a person has to create new choices for themselves. This is of particular importance, as every situation is unique and impossible to predict in detail. Creating new possibilities, and choosing between them, requires knowledge of the choices that also can form the basis for imagining novel choices.

Our discussion of how datanomous technology affects human autonomy emphasized that knowledge of the system and the data was necessary for the human to utilize and trust the technology, even if this meant facilitating the environment to be more technology-friendly. Not understanding the technology, or even knowing that the data eliminates choices, reduces the space for action for the human user.

Working with the concept revealed a need for more specificity about the differences in autonomy between humans and machines. The difference becomes visible when we examine their respective spaces for action. The space for action of autonomous technology is limited by the processing abilities of the technology and the data types that the data gathering equipment is able to register and process. In light of this, *datanomous technology* becomes a more precise concept. Human space for action is characterized by the ability to include fundamentally new types of data and actions into consideration, and to redefine a situation in profound ways. The incremental steps that a ML makes by adding new data to redefine its computing does not include the possibility of the kind of large and fundamental reconfigurations that a human can perform.

Datanomous technology can be part of the contextually situated resources that support human autonomy, but can also delimit human autonomy by reducing the human space for action. It is the purpose of the activity, instrumental or constitutive, that determines if automation strengthens or weakens human autonomy. If the purpose of the activity is an objective result, it is instrumental; if performing the activity is what gives it meaning, it is constitutive. An activity may change its purpose, and hence the human's role in it, if automation—or datanomos technology—is introduced.

Our exploration of the meanings of autonomy point to what is important for human autonomy: the type of activity and the space for action. In reference to responsible AI, it also points to what is important for design of datanomous technology *for* the autonomous human. Being aware of the central role of the data types for the technology, its limits and possibilities cease to be imperceptible. Thus, the human space for action can be expanded by *datanomous technology*.

Acknowledgments

The discussion in this paper originated in a workshop at NordiCHI 2020 conference. We thank all participants (Ralph Palmer, Patrick Björnfot, Julian Fietkau and Mandy Balthazar) for their ideas and good discussions during the workshop. We also thank our colleagues Svein Hovde, Jo Herstad, and Johanne Svanes Oskarsen for their contributions in the discussions before and after the workshop. We also wish to thank the participants in the two studies that we include in the paper (see notes 1 and 2). Both studies were carried out in line with the Norwegian and Swedish ethical boards, and all the participants gave written consent before participating in the research. Last, but not least, we wish to acknowledge that this work was partially supported by The Research Council of Norway (RCN) as a part of the projects: Multimodal Elderly Care Systems (MECS, Grant Agreement No. 247697), and Vulnerability in the Robot Society (VIROS, Grant Agreement No. 288285). Finally, we would like to thank the editors for critical comments and discussions of earlier versions of the paper.

Notes

1. The background for the robot vacuum cleaner example is a collection of empirical studies of robots in the wild, e.g., in non-lab settings, such as homes and gardens. In 2018-2020 we observed and interviewed 50 users of the robot vacuum cleaners iRobot Roomba and Neato from Neato Robotics about their experiences. These experiences were presented and discussed in a couple of studies, including critical reflections with regard to the ethical dilemmas of using such connected devices (see Saplacan et al., 2021; Saplacan, Herstad, & Pajalic, 2020b; Saplacan, Herstad, Tørresen, et al., 2020; Saplacan & Herstad, 2019). Our studies confirmed and supplemented other studies of everyday technologies in the home
2. The example of Lisa and her new insulin pump are part of a study presented in Schimmer (2021), where individuals with type 2 diabetes share their experience of living with diabetes.

References

Anderson, C. J. (2003). The psychology of doing nothing: Forms of decision avoidance result from reason and emotion. *Psychological Bulletin*, *129*(1), 139. https://doi.org/10.1037/0033-2909.129.1.139

Bainbridge, L. (1983). Ironies of automation. *Automatica*, *19*(6), 775-779. https://doi.org/10.1016/0005-1098(83)90046-8

Bianchini, S., Levillain, F., Menicacci, A., Quinz, E., & Zibetti, E. (2016). Towards Behavioral Objects: A Twofold Approach for a System of Notation to Design and Implement Behaviors in Non-anthropomorphic Robotic Artifacts. In J.-P. Laumond & N. Abe (Eds.), *Dance Notations and Robot Motion* (pp. 1-24). Springer International Publishing. https://doi.org/10.1007/978-3-319-25739-6_1

Bowker, G. C., & Star, S. L. (2008). *Sorting things out: Classification and its consequences* (1. paperback ed., 8. print). MIT Press.

Bradshaw, J. M., Hoffman, R. R., Woods, D. D., & Johnson, M. (2013). The Seven Deadly Myths of "Autonomous Systems." *IEEE Intelligent Systems*, *28*(3), 54-61. https://doi.org/10.1109/MIS.2013.70

Bratteteig, T. (2004). *Making Change. Dealing with relations between design and use* [Dr. Philos]. University of Oslo, Department of Informatics, Faculty of Mathematics and Natural Sciences.

Bratteteig, T., & Wagner, I. (2013). Moving Healthcare to the Home: The Work to Make Homecare Work. In O. W. Bertelsen, L. Ciolfi, M. A. Grasso, & G. A. Papadopoulos (Eds.), *ECSCW 2013: Proceedings of the 13th European Conference on Computer Supported Cooperative Work, 21-25 September 2013, Paphos, Cyprus* (pp. 143-162). Springer London. https://doi.org/10.1007/978-1-4471-5346-7_8

Brereton, M. (2013). Habituated objects: Everyday tangibles that foster the independent living of an elderly woman. *Interactions*, *20*(4), 20-24. https://doi.org/10.1145/2486227.2486233

Briefs, U., Ciborra, C., Schneider, L., & International Federation for Information Processing (Eds.). (1983). Systems design for, with, and by the users: Proceedings of the IFIP WG 9.1 Working Conference on Systems Design for, with, and by the Users, Riva del Sole, Italy, 20-24 September 1982. North-Holland Pub. Co. ; Sole distributors for the U.S.A. and Canada, Elsevier Science Pub. Co.

Broussard, M. (2018). Artificial Unintelligence: How Computers Misunderstand the World. MIT Press.

Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, *3*(1), 205395171562251. https://doi.org/10.1177/2053951715622512

Carr, N. G. (2015). The Glass Cage: How Our Computers Are Changing Us. Norton & Company.

Christiansen, E., & Andersen, P. V. K. (2013). Digital Living at Home—User Voices about Home Automation and a Home-Keeping Design Discourse. In M.

Aanestad & T. Bratteteig (Eds.), *Nordic Contributions in IS Research* (pp. 40-52). Springer. https://doi.org/10.1007/978-3-642-39832-2_3

Ciborra, C., & Willcocks, L. (2006). The Mind or the Heart? It Depends on the (definition of) Situation. *Journal of Information Technology*, *21*(3), 129-139. https://doi.org/10.1057/palgrave.jit.2000062

Code, L. (Ed.). (2004). *Encyclopedia of feminist theories* (Reprinted). Routledge.

Coeckelbergh, M. (2017). Language and technology: Maps, bridges, and pathways. *AI & SOCIETY*, *32*(2), 175-189. https://doi.org/10.1007/s00146-015-0604-9

Cohen, M. D., March, J. G., & Olsen, J. P. (1972). A Garbage Can Model of Organizational Choice. *Administrative Science Quarterly*, *17*(1), 1-25. https://doi.org/10.2307/2392088

Cummings, M. (2004, September 20). Automation Bias in Intelligent Time Critical Decision Support Systems. *AIAA 1st Intelligent Systems Technical Conference*. AIAA 1st Intelligent Systems Technical Conference, Chicago, Illinois. https://doi.org/10.2514/6.2004-6313

Declaration of Montréal for a responsible development of AI. (2018). Respaideclaration. https://www.montrealdeclaration-responsibleai.com/_files/ugd/ebc3a3_506ea08298cd4f8196635545a16b071d.pdf

Dignum, V. (2018). Ethics in artificial intelligence: Introduction to the special issue. *Ethics and Information Technology*, *20*(1), 1-3. https://doi.org/10.1007/s10676-018-9450-z

Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, *55*(10), 78-87. https://doi.org/10.1145/2347736.2347755

Dworkin, G. (1988). *The Theory and Practice of Autonomy*. Cambridge University Press. https://doi.org/10.1017/CBO9780511625206

Feinberg, M., Carter, D., Bullard, J., & Gursoy, A. (2017). Translating Texture: Design as Integration. *Proceedings of the 2017 Conference on Designing Interactive Systems*, 297-307. https://doi.org/10.1145/3064663.3064730

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds and Machines*, *28*(4), 689-707. https://doi.org/10.1007/s11023-018-9482-5

Forlizzi, J., & DiSalvo, C. (2006). Service robots in the domestic environment: A study of the roomba vacuum in the home. *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction*, 258-265. https://doi.org/10.1145/1121241.1121286

Formosa, P. (2021). Robot Autonomy vs. Human Autonomy: Social Robots, Artificial Intelligence (AI), and the Nature of Autonomy. *Minds and Machines*, 595-616 https://doi.org/10.1007/s11023-021-09579-2

Gendlin, E. T. (1978). Befindlichkeit: Heidegger and the philosophy of psychology. *Review of Existential Psychology and Psychiatry*, *16*(1-3), 43-71.

Gitelman, L. (Ed.). (2013). *"Raw Data" Is an Oxymoron*. MIT Press.

Goodwin, C. (1994). Professional Vision. *American Anthropologist*, *96*(3), 606-633. https://doi.org/10.1525/aa.1994.96.3.02a00100

Haapoja, J., & Lampinen, A. (2018). "Datafied" reading: Framing behavioral data and algorithmic news recommendations. *Proceedings of the 10th Nordic Conference on Human-Computer Interaction*, 125-136. https://doi.org/10.1145/3240167.3240194

Harper, R., Randall, D., & Sharrock, W. (2017). *Choice*. John Wiley & Sons.

Holmquist, L. E. (2017). Intelligence on tap: Artificial intelligence as a new design material. *Interactions*, *24*(4), 28-33. https://doi.org/10.1145/3085571

Hutchins, E. (2006). *Cognition in the wild* (8. pr). MIT Press.

Ju, W., & Leifer, L. (2008). The Design of Implicit Interactions: Making Interactive Systems Less Obnoxious. *Design Issues*, *24*(3), 72-84. https://doi.org/10.1162/desi.2008.24.3.72

Kahneman, D. (2011). *Thinking, fast and slow* (1st ed). Farrar, Straus and Giroux.

Langefors, B. (1980). Infological models and information user views. *Information Systems*, *5*(1), 17-32. https://doi.org/10.1016/0306-4379(80)90065-4

Luck, M., D'Inverno, M., & Munroe, S. (2003). Autonomy: Variable and Generative. In H. Hexmoor, C. Castelfranchi, & R. Falcone (Eds.), *Agent Autonomy* (pp. 11-28). Springer US. https://doi.org/10.1007/978-1-4419-9198-0_2

Mackenzie, C., & Stoljar, N. (2000). *Relational autonomy: Feminist perspectives on autonomy, agency, and the social self.* Oxford University Press. https://researchers.mq.edu.au/en/publications/relational-autonomy-feminist-perspectives-on-autonomy-agency-and-

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955. *AI Magazine*, *27*(4), 12-12. https://doi.org/10.1609/aimag.v27i4.1904

Mikalef, P., Conboy, K., Lundström, J. E., & Popovič, A. (2022). Thinking responsibly about responsible AI and 'the dark side' of AI. *European Journal of Information Systems*, *0*(0), 1-12. https://doi.org/10.1080/0960085X.2022.2026621

Mills, C. W. (1967). *The sociological imagination* (1. iss. as an Oxford Univ. Press paperb., 43. print). Oxford Univ. Press.

Mol, A. (2008). The logic of care: Health and the problem of patient choice. Routledge.

Muller, M., Lange, I., Wang, D., Piorkowski, D., Tsay, J., Liao, Q. V., Dugan, C., & Erickson, T. (2019). How Data Science Workers Work with Data: Discovery, Capture, Curation, Design, Creation. *Proceedings of the 2019*

CHI Conference on Human Factors in Computing Systems, 1-15. https://doi.org/10.1145/3290605.3300356

Nygaard, K. (1996). "Those Were the Days"? Or "Heroic Times Are Here Again"? *Scandinavian Journal of Information Systems*, *8*(2). https://aisel.aisnet.org/sjis/vol8/iss2/6

Nyholm, S. (2018). The ethics of crashes with self-driving cars: A roadmap, I. *Philosophy Compass*, *13*(7), e12507. https://doi.org/10.1111/phc3.12507

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans*, *30*(3), 286-297. https://doi.org/10.1109/3468.844354

Pearsall, J. (Ed.). (1999). *The concise Oxford dictionary* (10. ed). Oxford Univ. Press.

Pine, K. H., & Liboiron, M. (2015). The Politics of Measurement and Action. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 3147-3156. https://doi.org/10.1145/2702123.2702298

Purves, D., Jenkins, R., & Strawser, B. J. (2015). Autonomous Machines, Moral Judgment, and Acting for the Right Reasons. *Ethical Theory and Moral Practice*, *18*(4), 851-872. https://doi.org/10.1007/s10677-015-9563-y

Russell, S. J., & Norvig, P. (1995). *Artificial Intelligence: A Modern Approach*. Prentice Hall.

SAE. (n.d.). SAE Levels of Driving AutomationTM Refined for Clarity and International Audience. Retrieved February 25, 2022, from https://www.sae.org/site/blog/sae-j3016-update

Säljö, R. (2001). Læring i praksis et sosiokulturelt perspektiv. Cappelen akademisk.

Saplacan, D. (2020). *Situated Abilities: Understanding Everyday Use of ICTs* [Ph.D. Thesis, University of Oslo, Department of Informatics, Faculty of Mathematics and Natural Sciences]. https://www.duo.uio.no/handle/10852/81852

Saplacan, D., & Herstad, J. (2019). An Explorative Study on Motion as Feedback: Using Semi-Autonomous Robots in Domestic Settings. *International Journal On Advances in Software*, *12*(1 and 2), 68-90.

Saplacan, D., Herstad, J., & Pajalic, Z. (2020). An Analysis of Independent Living Elderly's Views on Robots: A Descriptive Study from the Norwegian Context. *Proceedings of The Thirteenth International Conference on Advances in Computer-Human Interactions*, 199-208. http://www.thinkmind.org/index.php?view=article&articleid=achi_2020_5_10_20011

Saplacan, D., Herstad, J., Tørresen, J., & Pajalic, Z. (2020). A Framework on Division of Work Tasks between Humans and Robots in the Home. *Multimodal Technologies and Interaction*, *4*(3), 44. https://doi.org/10.3390/mti4030044

Saplacan, D., Khaksar, W., & Torresen, J. (2021). On Ethical Challenges Raised by Care Robots: A Review of the Existing Regulatory-, Theoretical-, and Research Gaps. *2021 IEEE International Conference on Advanced Robotics and Its Social Impacts (ARSO)*, 219-226. https://doi.org/10.1109/ARSO51874.2021.9542844

Schmidt, C. T. A., & Kraemer, F. (2006). Robots, Dennett and the autonomous: A terminological investigation. *Minds and Machines*, *16*(1), 73-80. https://doi.org/10.1007/s11023-006-9014-6

Schneewind, J. B. (1998). The invention of autonomy: A history of modern moral philosophy. Cambridge University Press.

Schütz, A. (1951). Choosing Among Projects of Action. *Philosophy and Phenomenological Research*, *12*(2), 161-184. https://doi.org/10.2307/2103478

Simonsen, J., & Robertson, T. (Eds.). (2013). *Routledge international handbook of participatory design*. Routledge Taylor & Francis.

Soltanzadeh, S. (2021). Strictly Human: Limitations of Autonomous Systems. *Minds and Machines*. https://doi.org/10.1007/s11023-021-09582-7

Soma, R. (2020). Engaging in Deep Wonder at the Experience of Encountering a Lawnmower Robot. In M. Nørskov, J. Seibt, & O. S. Quick (Eds.), *Frontiers* *in Artificial Intelligence and Applications*. IOS Press. https://doi.org/10.3233/FAIA200905

Soma, R., & Herstad, J. (2018). Turning Away from an Anthropocentric View on Robotics. *Frontiers in Artificial Intelligence and Applications*, 53-62. https://doi.org/10.3233/978-1-61499-931-7-53

Soma, R., Søyseth, V. D., Søyland, M., & Schulz, T. W. (2018). *Facilitating Robots at Home: A Framework for Understanding Robot Facilitation*. International Academy, Research and Industry Association (IARIA).

Suchman, L. A. (1987). Plans and situated actions: The problem of human-machine communication. Cambridge University Press.

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness* (pp. x, 293). Yale University Press.

Vassilakopoulou, P., Grisot, M., Parmiggiani, E., & Shollo, A. (n.d.). *Responsible AI—Critical Perspectives on AI-infused Digitalization*. Scandinavian Journal. http://iris.cs.aau.dk/index.php/special-issue-call-for-papers.html

Verne, G. (2014). Two faces of autonomy: Learning from non-users of an e-service. Systems, Signs & Actions—An International Journal on Information Technology, Action, Communication and Workpractices, 8(1), 6-24.

Verne, G. (2020). Adapting to a Robot: Adapting Gardening and the Garden to fit a Robot Lawn Mower. *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, 34-42. https://doi.org/10.1145/3371382.3380738

Weizenbaum, J. (1976). Computer power and human reason: From judgment to calculation (pp. xii, 300). W. H. Freeman & Co.

Winograd, T., & Flores, F. (1986). Understanding Computers and Cognition: A New Foundation for Design. Intellect Books.

Xu, W. (2021). From Automation to Autonomy and Autonomous Vehicles: Challenges and Opportunities for Human-Computer Interaction. *Interactions*, *28*. https://doi.org/10.1145/3434580