IoT

Article

Achieving Ethical Algorithmic Behaviour in the Internet of Things: A Review

Seng W. Loke

School of Information Technology, Deakin University, Geelong 3217, Australia; seng.loke@deakin.edu.au

Abstract: The Internet of Things is emerging as a vast, inter-connected space of devices and things surrounding people, many of which are increasingly capable of autonomous action, from automatically sending data to cloud servers for analysis, changing the behaviour of smart objects, to changing the physical environment. A wide range of ethical concerns has arisen in their usage and development in recent years. Such concerns are exacerbated by the increasing autonomy given to connected things. This paper reviews, via examples, the landscape of ethical issues, and some recent approaches to address these issues concerning connected things behaving autonomously as part of the Internet of Things. We consider ethical issues in relation to device operations and accompanying algorithms. Examples of concerns include unsecured consumer devices, data collection with health-related Internet of Things, hackable vehicles, behaviour of autonomous vehicles in dilemma situations, accountability with Internet of Things systems, algorithmic bias, uncontrolled cooperation among things, and automation affecting user choice and control. Current ideas towards addressing a range of ethical concerns are reviewed and compared, including programming ethical behaviour, white-box algorithms, black-box validation, algorithmic social contracts, enveloping IoT systems, and guidelines and code of ethics for IoT developers; a suggestion from the analysis is that a multi-pronged approach could be useful based on the context of operation and deployment.

Keywords: ethics for the Internet of Things; ethical algorithmic behaviour; ethics

1. Introduction

The Internet of Things (or IoT, for short) involves devices or things connected to the internet or with networking capability. This includes internet devices, such as smartphones, smartwatches, smart TVs, smart appliances, smart cars, and smart drones, as well as everyday objects with Bluetooth, 3G/4G, and Wi-Fi capabilities. Specialised IoT protocols, such as NB-IoT, Sigfox and LoRAWAN, provide new connectivity options for the IoT (https://www.rs-online.com/designspark/eleven-internet-of-things-iot-protocols-you-need-to-know-about, accessed on 1 July 2021).

Apart from industrial IoT systems, what is beginning to emerge is the notion of everyday objects with the following:

- Internet or network connectivity (e.g., WiFi or Bluetooth enabled).
- Sensors (e.g., think of the sensors in the smartphone but also in a fork to detect its movement and how fast people eat [1,2] (See https://www.hapilabs.com/product/hapifork, accessed on 1 July 2021)).
- Computational ability (e.g., with embedded AI [3] and cooperation protocols).
- Actuators, or the ability to affect the physical world, including taking action autonomously.

The above highlights only some aspects of the IoT—an extensive discussion on the definition of the Internet of Things is in [4].

There are also new home appliances, such as Amazon Alexa (https://developer.amazon.com/alexa, accessed on 1 July 2021) and Google Home (https://madeby.google.com/ho

me/, accessed on 1 July 2021), which have emerged with internet connectivity being central to their functioning. Often, they can be used to control other devices in the home. When things are not only internet-connected, but addressable via Web links or URLs (the Uniform Resource Locators), and communicate via Web protocols (e.g., using the Hypertext Transfer Protocol (HTTP)) the so-called Web of Things (https://www.w3.org/WoT/, accessed on 1 July 2021) emerges. An example of early work on the social Web of Things is the Paraimpu platform [5].

With increasing autonomy (fuelled by developments in AI) and connectivity (fuelled by developments in wireless networking), there are a number of implications:

- Greater cooperation among IoT devices can now happen; devices that were previously not connected can now not only communicate (provided that the time and resource constraints allow), but also carry out cooperative behaviours. In fact, the work in [6] envisions universal machine-to-machine collaboration across manufacturers and industries by the year 2025, though this can be restricted due to proprietary data. Having a cooperation layer above the networking layer is an important development; the social IoT has been widely discussed [7–10];

- Network effects emerge. The value of a network is dependent on the size of the network; the greater the size of the network, the greater the value of joining or connecting to the network, so device manufacturers may tend to favour cooperative IoT (e.g., see the Economics of the Internet of Things (https://www.technologyreview.com/s/527361/the-economics-of-the-internet-of-things/, accessed on 1 July 2021)). A device that can cooperate with more devices could have greater value, compared to ones that cooperate with only a few—such cooperation among devices can be triggered by users directly or indirectly (if decided by a device autonomously), with a consequent impact on communication latency and delay.

- Devices that are connected to the internet are controllable via the internet, which means they are also vulnerable to (remote) hacking, in the same way that a computer being on the internet can be hacked.

- Sensors on such IoT devices gather significant amounts of data and, being internet-enabled, such data are typically uploaded to a server (or to a Cloud computing server somewhere); potentially, such data can cause issues with people who are privacy-conscious (e.g., data from an internet-connected light bulb could indicate when someone is home and not home). An often-linked topic to the IoT is the notion of data analytics, due to the need to process and analyse data from such sensing devices. There are also ethical and privacy issues of placing sensors in different areas (e.g., in certain public areas) and there could be cultural sensitivities in relation to where sensors are placed.

- IoT devices may be deployed over a long time (e.g., embedded in a building or as part of urban street lighting) so they need to be upgraded (or their software upgraded) over the internet as improvements are made, errors are found and fixed, and as security vulnerabilities are discovered and patched.

- Non-tech savvy users might find working with internet-connected devices challenging (e.g., set up and maintenance, and be unaware of security or privacy effects of devices), and the users might feel a loss of control.

- Computation on such devices suggests that greater autonomy and more complex decision-making is possible (and devices with spare capacity can also be used to supplement other devices); in fact, autonomous behaviour in smart things are not new. Smart things detecting sensor-based context and responding autonomously (using approaches ranging from simple Even-Condition-Action rule-based approaches to sophisticated reasoning in agent-based approaches) have been explored extensively in context-aware computing [11,12].

From the above, one can see that the IoT offers tremendous opportunity, but also raises a range of ethical concerns. Prominent computer scientists have noted the need for ethical policies to guide IoT governance in the areas of privacy rights, accountability for

autonomous systems, and promoting the ethical use of technologies [13]. Ethical concerns regarding the use of IoT devices as well as with datafication and algorithmisation in the context of smart cities have also been raised [14,15].

This paper aims to review the landscape of ethical concerns and issues that have arisen and which could, in future, arise with the Internet of Things, focusing on device operations and accompanying algorithms, especially as algorithms are increasingly embedded in, and run on, IoT devices, enabling devices to take action with increasing autonomy; the paper also reviews current ideas for addressing these concerns. Based on the review, it is suggested that a multi-pronged approach can be useful for achieving ethical algorithmic behaviour in connected things.

1.1. Scope and Context

There has been much recent thinking on how artificial intelligence (AI) technologies can be integrated with IoT, from applying AI algorithms to learn from IoT data, multiagent views of IoT, to connected robots [3,16] (https://emerj.com/ai-sector-overviews/artificial-intelligence-plus-the-internet-of-things-iot-3-examples-worth-learning-from/, accessed on 1 July 2021). As AI capabilities become embedded in IoT devices, the devices gain greater autonomy and decision-making capabilities, automating a wider range of tasks, so much so that some things can be described as "robotic". For example, we can imagine a bookshelf that one can talk to and which can serve us, relocating and reorganising books at our command; a library where the storage and retrieval of (physical) books is automated; or think of a standing lamp that follows and tracks the user as the user moves around on the sofa or in the room. The question is whether these libraries and the standing lamp can be considered "robots". Additionally, autonomous connected vehicles [17], with internet-enabled networking and vehicle-to-vehicle connectivity have also captured the world's imagination and have enjoyed tremendous developments in recent years. The discussion in this paper, therefore, includes robots, AI used with IoT, and autonomous vehicles, under the broad definition of IoT. The link between IoT and robotics was also made in [18,19], yielding the notion of the Internet of Robotic Things.

Ethics in AI are extensively discussed elsewhere (e.g., [20] and ethical AI principles (https://futureoflife.org/ai-principles/, accessed on 1 July 2021)), and indeed, the integration of AI technologies into the IoT, as mentioned above, calls for ethical AI principles to be considered in the development of such IoT systems. Hence, this paper reviews work not only for ethical issues in IoT, but also includes a review of work on ethics in AI and robotics in the context of IoT.

While this paper reviews technical approaches mainly from computing, the issues are often interdisciplinary, at the intersection of computing technology, philosophy (namely, ethics), law and governance (insofar as policies are involved), as well as diverse application domains (e.g., health, transport, business, defence, law, and others) where IoT has been applied. Moreover, while security and data privacy are key concerns in relation to things behaving ethically, the concerns on ethical behaviour go beyond security and privacy. The field is also continually growing to date, as ethical issues for IoT are highlighted in mass media and with much research in the area (examples highlighted in the following sections); hence, the paper does not seek to completely cover all recent work, but can only provide a comprehensive snapshot and introduction to the area, while highlighting potential approaches to the issues.

The seminal review on the ethics of computing [21] lays out five aspects of each paper reviewed: ethical theory that aids interpreting the issue, the methodology applied, the technology context, and contributions and recommendations. Different from [21], this paper focuses on the ethical issues in IoT work, but, indeed, these aspects inform the reading of work in this area at the junction of the IoT and ethics. This paper touches on a range of ethical issues noted in the paper, namely, agency, autonomy, consent, health, privacy, professionalism, and trust, in our discussions. For example, we discuss issues of user choice and consent in IoT devices, the autonomy of things in their function, health

IoT issues, the security, trust and privacy of IoT devices, and the code of ethics for IoT developers. We do not discuss ethical issues in relation to inclusion and the digital divide but retain a technical focus in this paper.

The survey on foundational ethical issues in IoT [22] focused on informed consent, privacy, security, physical safety, and trust and noted, importantly, that these are inter-related in IoT. This paper also discusses a range of these issues but we also consider examples and solutions (many originally from outside typical IoT research areas) to achieving ethical IoT systems. Ethical and trust issues with IoT are also discussed in [23], among other concerns, such as IoT laws and smart contracts.

1.2. Organization

The rest of the paper is organised as follows. To introduce readers to ethical issues in IoT, Section 2 first discusses, via examples, ethical concerns with IoT. Then, Section 3 examines ideas that have been proposed to address these concerns, and notes the need for a multi-pronged approach. Section 4 concludes with future work.

2. Ethical Concerns and Issues

This section reviews ethical concerns and issues with IoT devices and systems, via examples in multiple application domains, including the need for consumer IoT devices to employ adequate security measures, ethical data handling by health related IoT systems, right behaviour of autonomous vehicles in normal usage and dilemma situations, usage concerns with connected robots and ethical robot behaviour, algorithmic bias that could be embedded into IoT systems, right behaviour when IoT devices cooperate, and user choice restrictions or lost of control with automated IoT systems. Below, the unit of analysis is either an individual IoT device or a collection (or system) of such IoT devices (the size of which depends on the application).

2.1. Unsecured Consumer IoT Devices

The security and data privacy issues in IoT are well surveyed and have been discussed extensively, e.g., in [24–33]. The contents of the surveys are not repeated here but some examples of issues with unsecured IoT devices are highlighted below.

There are IoT devices which may have been shipped without encryption (with lower computational power devices which are not capable of encrypted communications). A study by HP (https://www.hp.com/us-en/hp-news/press-release.html?id=1744676 #.YOAh2C0RqK4, accessed on 1 July 2021) suggested that 70 percent of IoT devices use unencrypted communications. However, it must be noted that cheaper does not necessarily mean less secure, as cost depends on a range of factors beyond security capability.

A Samsung TV was said to listen in on living room conversations as it picks up commands via voice recognition. The company has since clarified that it does not record conversations arbitrarily (http://abcnews.go.com/Technology/samsung-clarifies-privacy-policy-smart-tv-hear/story?id=28861189, accessed on 1 July 2021). However, it does a raise a concern with devices in the same category as voice-activated or conversational devices as to whether or not they do record conversations.

In an experiment (https://www.timesofisrael.com/israeli-hackers-show-light-bulbs-can-take-down-the-internet/, accessed on 1 July 2021) at Israel's Weizmann Institute of Science, researchers managed to fly a drone within 100 metres of a building and remotely infect light bulbs in the building by exploiting a weakness in the ZigBee Light Link protocol, used for connecting to the bulbs. The infected bulbs were then remotely controlled via the drone and made to flash 'SOS'.

A report on the Wi-Fi enabled Barbie doll (https://www.theguardian.com/techno logy/2015/nov/26/hackers-can-hijack-wi-fi-hello-barbie-to-spy-on-your-children, accessed on 1 July 2021) noted that they can be hacked and turned into surveillance devices. This was then followed by an FBI advisory note on IoT toys, (https://www.ic3.gov/media/2017/170717.aspx, accessed on 1 July 2021) about possible risk to private

information disclosure. An 11 year old managed to hack into a teddy bear via Bluetooth (https://securityintelligence.com/news/with-teddy-bear-bluetooth-hack-11-year-old-proves-iot-security-is-no-childs-play/, accessed on 1 July 2021). Additionally, ubiquitous IoT cameras have certainly not been free from hacking (http://www.zdnet.com/article/175000-iot-cameras-can-be-remotely-hacked-thanks-to-flaw-says-security-researcher/, accessed on 1 July 2021).

There are many other examples of IoT devices getting hacked (https://www.wired.com/2015/12/2015-the-year-the-internet-of-things-got-hacked/, accessed on 1 July 2021) As research shows, (https://arxiv.org/pdf/1705.06805.pdf, accessed on 1 July 2021) someone can still infer people's private in-home activities by monitoring and analysing network traffic rates and packet headers, even when devices use encrypted communications.

The above are only a few of many examples and has implications for developers of IoT devices, which must incorporate security features, for policy-makers, for cyber security experts, as well as for users who need to be aware of potential risks.

Recent surveys also highlighted privacy and managerial issues with the IoT [34]. From the Australian privacy policy perspective [35], after a review of the four issues of (1) IoT-based surveillance, (2) data generation and use, (3) inadequate authentication and (4) information security risks, the conclusion is that the Australian Privacy Principle is inadequate for protecting the individual privacy of data collected using IoT devices and that, given the global reach of IoT devices, privacy protection legislation is required across international borders. Weber [36] calls for new legal approaches to data privacy in the IoT context from the European perspective, based on improved transparency and data minimization principles. The recent European regulation, the General Data Protection Regulation (GDPR) (https://gdpr.eu, accessed on 1 July 2021) is a law aimed at providing people with greater control of their data and has implications and challenges for IoT systems, with requirements on systems, such as privacy-by-design, the right to be forgotten or data erasure, the need for clarity in requesting consent, and data portability, where users have the right to receive their own data as discussed in [37]. Companies are already coming on board with tools to support GDPR requirements (for example, see Microsoft's tools: https://www.microsoft.com/en-au/trust-center/privacy/gdpr-overview, accessed on 1 July 2021 and Google: https://cloud.google.com/security/compliance/gdpr/, accessed on 1 July 2021).

A recent workshop on privacy and security policies for IoT at Princeton University (https://spia.princeton.edu/events/conference-security-and-privacy-internet-things-iot, accessed on 1 July 2021) raised a range of issues, suggesting a still on-going area of research at the intersection of IT, ethics, governance and law. Cyber physical systems security is discussed extensively elsewhere [38].

Security also impacts usability since additional measures might be taken to improve security, for example, when users have to reset passwords before being allowed to use a device, the use of multi-factor authentication schemes, and configuration set up to improve security during use, all of which could reduce usability; the work in [39] highlights the need to consider the usability impact of IoT security features at the time of design.

2.2. Ethical Issues with Health Related IoT

IoT medical devices play an increasingly critical role in human life, but as far back as 2008, implantable defibrillators were known to be 'hackable' [40], allowing communications to them to be intercepted.

Apart from the security of IoT devices, in [41], a range of ethical issues with the use of IoT in health were surveyed, including the following:

- *Personal privacy*: This relates not just to privacy of collected data, but the notion that a person has the freedom not to be observed or to have his/her own personal space. The use of smart space monitoring (e.g., a smart home or in public spaces such as aged care facilities) of its inhabitants raises the concern of continual observation of

people, even if it is for their own good—being able to monitor individuals or groups can be substantially beneficial but presents issues of privacy and access.

- *Informational privacy*: This relates to one's ability to control one's own health data. It is not uncommon for organizations to ask consumers for private data with the promise that the data will not be misused—in fact, privacy laws can prohibit the use of data beyond its intended context. The issues are myriad (e.g., see [42]), including how one can access data that were collected by an IoT device but are now possibly owned by the company, how much an insurance company can demand of user health data, (https://www.iothub.com.au/news/intel-brings-iot-to-health-insurance-411714, accessed on 1 July 2021), how one can share data in a controlled manner, how one can prove the veracity of personal health data, and how users can understand the privacy-utility trade-offs when using an IoT device.

- *Risk of non-professional care*: The notion of self health-monitoring and self-care as facilitated by health IoT devices can provide false optimism, limiting a patient's condition to a narrow range of device-measurable conditions. Confidence in non-professional carers armed with IoT devices might be misplaced.

The above issues relate mainly to health IoT devices, but the data privacy issues apply to other internet-connected devices in general [43] (see also http://arno.uvt.nl/show.cgi?fid=144871, accessed on 1 July 2021). Approaches to data privacy range from privacy-by-design, recent blockchain-based data management and control (e.g., [44–48]), to regulatory frameworks that aim to provide greater control over data for users as reviewed earlier, e.g., in [26,32,33,37]. There are also issues related to how health data should or should not be used—for example, what companies are allowed to use the health data for (e.g., whether an individual could be denied insurance based on health data, or denied access to treatment based on lifestyle monitoring).

In relation to IoT in sports to help sports training and fitness tracking, incorporated in an artificial personal trainer, there are numerous technical challenges [49], including generating and adapting plans for the person, measuring the person's readiness, personal data management, as well as validation and verification of fitness data. One could think of issues and liability with errors in measurement or an athlete being endangered or, subsequently, even injured over time by erroneous advice due to incorrect measurements, or issues arising from following the advice of an AI trainer or such devices being hacked. In any case, there are already several wearable personal trainers on the market (e.g., see https://welcome.moov.cc/ and https://vitrainer.com/pages/vi-sense-audio-trainer, accessed on 1 July 2021), which come with appropriate precautions and disclaimers for users (See https://welcome.moov.cc/terms/, accessed on 1 July 2021 and https://vitrainer.com/pages/terms-and-conditions) and privacy policies, accessed on 1 July 2021.

2.3. Hackable Vehicles and the Moral Dilemma for Autonomous Vehicles (AVs)

Cars with computers are not unhackable; an example is the Jeep, which was hacked while on the road (https://www.wired.com/2015/07/hackers-remotely-kill-jeep-highway/, accessed on 1 July 2021) and made to be remotely controllable. With many vehicles having internet connectivity, their hackability is now public knowledge (for example, see the online book on car hacking, http://opengarages.org/handbook/, accessed on 1 July 2021). Similar security issues of encrypting communications with the vehicle and securing the vehicle system arise, as well as those for managing data collected by the vehicle, as in other IoT systems—security issues for autonomous vehicles are discussed elsewhere [50,51]. Given the wide range of data collected about vehicles, from telemetry data to location data, as well as logs of user interaction with the vehicle computer, privacy management of vehicular data is an issue [52].

Recent developments in autonomous vehicles have provided tremendous promise for reducing road injuries and deaths due to human error, as well as the potential to start a 'new' industry, with many countries around the world working on autonomous vehicle projects, with a subsequent impact on the design of cities. However, as autonomous

vehicles function in a socio-technical environment, there may be decisions that they need to make, which involve moral reasoning as discussed in [53] (see also this TED talk by https://rahwan.me/Professor Iyad Rahwan from MIT, accessed on 1 July 2021).

Essentially, the moral dilemma of autonomous vehicles is similar to the trolley problem in philosophy (see https://fee.org/articles/the-trolley-problem-and-self-driving-cars/, accessed on 1 July 2021)—suppose that an autonomous vehicle is about to hit someone in its way and the only way to avoid this is to swerve to the right or left, but it will kill some pedestrians while doing so. Should it opt to protect the occupants of the vehicle in preference to external individuals? As, either way, someone will be killed, what should the autonomous vehicle do? (Such dilemma situations can occur in other smart things scenarios—e.g., consider this original example: in a fire situation, a smart door can decide to open to let someone through but this, at the same time, would allow smoke to pass in to possibly harm others. Or, a smart thing can choose to transmit messages about a lost dementia-afflicted person frequently to allow finer-grained tracking for a short time, but risk the battery running out sooner (and so lose the person if not found in time), or transmit less frequently allowing a longer operating time but coarse-grained location data.) While there may be no clear-cut answer to the question, it is important to note the ethical issue raised—potential approaches to the problem are discussed later. While AVs will help many people, there are issues about what the AVs will do in situations where trade-offs are required. A utilitarian approach might be to choose the decision that potentially kills fewer people. A virtue ethics approach will not approve of that way of reasoning. A deontological or virtuous approach might decide 'not to kill', whatever the situation, in which case the situation cannot be helped. One could also argue that such situations are unlikely to arise, but there is a small possibility that it could arise and perhaps in many different ways. Imagine an AV in a busy urban area receiving instruction to speed up due to a heart attack that just occurred in its passenger, putting other pedestrians and road users at greater risk—should the AV speed up? However, one could also note that sensors in the vehicle could detect that the passenger has a heart attack and report this to traffic management to have a path cleared, and so, speeding up may not be an issue—hence, connectivity can help the situation rather than increase risk, while the ethics in the decision making remains challenging.

Ethical guidelines regulating the use and development of AVs are being developed—Germany was the first country to provide ethical guidelines for autonomous vehicles via the Ethics Commission on Connected and Automated Driving (see the report at https://www.bmvi.de/SharedDocs/EN/Documents/G/ethic-commission-report.pdf?__blob=publicationFile, accessed on 1 July 2021). The guidelines include an admission that autonomous driving cannot be completely safe: "... *at the level of what is technologically possible today, and given the realities of heterogeneous and non-interlinked road traffic, it will not be possible to prevent accidents completely. This makes it essential that decisions be taken when programming the software of highly and fully automated driving systems.*" As noted in [54], for " *future autonomous cars, crash-avoidance features alone won't be enough*", but when crash is inevitable, there needs to be a crash-optimization strategy. However, that strategy should aim not only to minimise damage since if that was the case, the vehicle might decide to crash into a cyclist wearing a helmet than a cyclist not wearing a helmet, thus, targeting people who chose to be safer—there is no clear resolution to this ethical issue, as of yet.

There are also issues concerning who will take responsibility when something bad happens in an autonomous vehicle—whether it would be the passengers, the manufacturer or the middlemen. The issue is complex in mixed scenarios where human drivers and autonomous vehicles meet in an incident, and the fault lies with the human driver but the autonomous vehicle was unable to react to the human error.

Assuming the success of autonomous vehicles to reduce road deaths and accidents, would it then be ethical to allow human drivers? The work in [55] goes as far as to suggest "...*making it illegal to manufacture vehicles that allow for manual driving beyond a certain date and/or making it illegal, while on a public road, to manually drive a vehicle that has the capacity*

for autonomous operations." Appropriate policies for autonomous vehicles continues to be an open issue [56]. Further approaches to ethical automated vehicles are discussed in Section 3.

2.4. Roboethics

Roboethics [57–59] is concerned with the positive and negative aspects of robotics technology in society, and explores issues concerning the ethical design, development and use of robots. While there are tremendous opportunities in robotics, their widespread use also raises ethical concerns, and as the line between robots and autonomous IoT becomes blurred, the issues of ethics with robots are inherited by IoT.

2.4.1. Robots Rights

There are some schools of thought that have begun to ask the question of whether robots (if capable of moral reasoning) should have rights [60], and what level of autonomous decision making would require robots to have rights, similar to how animals might have rights. The issue of robot rights and robot interests from a legal perspective was raised in [61].

The level of autonomy and sentience required of such machines before rights becomes an issue might still be far off. In fact, roboethics has largely been concerned with the ethics that developers and users of the technology need to consider.

Below, we explore examples of ethical issues in robotics for surgery, personal assistance, and war.

2.4.2. Robotic Surgery

Robots are capable of surgical operations as we have seen, typically under the direction and control of a surgeon. In 2000, the U.S. Food and Drug Administration (FDA) approved the use of the Da Vinci robotic surgical system for a surgeon to perform laparoscopic gall bladder and reflux disease surgery. Robotic surgery devices continue to be developed (https://spectrum.ieee.org/robotics/medical-robots/would-you-trust-a-robot-surgeon-to-operate-on-you, accessed on 1 July 2021), and some make decisions autonomously during surgeries, e.g., automatically positioning a frame for the surgeon's tools, locating where to cut bones, and delivering radiation for tumours. If the costs of robotic surgery can be reduced, complex surgery could perhaps be made available to more people in third-world and developing countries. As they get better, and are able to provide surgical help at lower costs, what is problematic then is not their use, but denying people their use.

An issue emerges when something goes wrong and the question of accountability and liability arises regarding the patient's injury. While one might not consider surgical robots as IoT devices, the issue of IoT devices making decisions that could result directly in injury or harm, even if they were intentionally made to help humans, raises similar concerns.

2.4.3. Social and Assistive Robots and Smart Things

Social robots might play the role of avatars (remotely representing someone), social partners (accompanying someone at home), or cyborg extensions (being linked to the human body in some way). A robot capable of social interaction might be expected to express and perceive emotions, converse with users, imitate users, and establish social connection with users via gesture, gaze or some form of natural interaction modality, as well as, perhaps, present a distinctive personality. While they can be useful, some concerns include the following:

- Social robots or IoT devices may be able to form bonds with humans, e.g., an elderly person or a child. A range of questions arises, such as whether such robots should be providing emotional support in place of humans, if they can be designed to do so. Another question regards the psychological and physical risk of humans forming such bonds with such devices or robots—when a user is emotionally attached to a thing,

a concern is what would happen if the thing is damaged or no longer supported by the manufacturer, or if such things can be hacked to deceive the user. This question can be considered for smart things that have learnt and adapted to the person's behaviour and are not easily replaced.

- Such social robots or IoT devices can be designed to have authority to provide reminders, therapy or rehabilitation to users. Ethical issues can arise when harm or injury is caused due to interaction with such robots. For example, death caused from medication taken at the wrong time, due to a robot's reminder at the wrong time due to malfunction. A similar concern carries over to a smart pill bottle (an IoT device) intended to track when a person has and has not taken medication with an associated reminder system. There is also a question of harm being caused inadvertently, e.g., when an elderly person trips over a robot that approached too suddenly, or a robot makes decisions on behalf of its owner, without the owner's full consent or before the owner can intervene.

It must be noted, though, that the concerns above relate to behavioural aspects of the devices, not so much to the connectivity that the devices might have.

Ethical guidelines regarding their development and use are required, including the training of users and caregivers, the affordability of such devices, and the prevention of malpractice or misuse.

Ethical principles for socially assistive robots are outlined in a report (https://robotics .usc.edu/publications/media/uploads/pubs/689.pdf, accessed on 1 July 2021), including the following:

> *"The principles of beneficence and non-maleficence state that caregivers should act in the best interests of the patient and should do nothing rather than take any action that may harm a patient."*

A similar guideline informs socially-assistive smart things, not only robots—the question of how smart things with intelligent and responsive behaviours and robots could be programmed to provably satisfy those principles is an open research issue. It remains an open research issue as to how smart things and robots can learn human values and be flexible enough to act in a context-aware manner. Issues specifically due to the fact that these devices might be connected are similar to other IoT devices, e.g., sensitive private data possibly shared beyond safe boundaries, and vulnerability to remote hacking—perhaps made worse by their close interaction with and proximity to users.

2.4.4. Robots in War

Robots can be used to disarm explosive devices, for 24/7 monitoring, or for risky spying missions, and to engage in battles in order to save lives. However, there are already controversies surrounding the use of automated drones (even if remotely human piloted) in war [62]. While human casualties can be reduced, the notion of humans being out of the loop in *firing decisions* is somewhat controversial. AI also might not have adequate discriminatory powers for its computer vision technology to differentiate civilians from soldiers. While robots can reduce human lives lost at war, there is also the issue that it could then lower barriers to entry and even 'encourage' war, or be misused by 'tyrants'. There has been a call for a ban on autonomous weapons in an open letter signed by 116 founders of robotics and AI companies from 26 countries, (https://futureoflife.org/2017/08/20/killer-robots-worlds-top-ai-robotics-compa nies-urge-united-nations-ban-lethal-autonomous-weapons/, accessed on 1 July 2021) and the *Campaign to stop Killer Robots* (http://www.stopkillerrobots.org/, accessed on 1 July 2021). Algorithmic behaviour can be employed in remotely controlled robots to help human operators, but remote controlled and autonomous robotic weapons, if allowed, will need to be designed for responsibility, i.e., allow human oversight and control (https://www.oxfordmartin.ox.ac.uk/downloads/briefings/Robo-Wars.pdf, accessed on 1 July 2021). Robot-on-robot warfare might still be legal and ethical.

2.5. Algorithmic Bias and IoT

We explore the notion of bias in algorithms in this section. The following types of concerns with algorithms were noted by [63]: *inconclusive evidence* when algorithms make decisions based on approximations, machine-learning techniques, or statistical inference; *inscrutable evidence,* where the workings of the algorithm are not known; and *misguided evidence* when the inputs to the algorithm are inappropriate. Some automated systems have behaviour that can be opaque, unregulated and could amplify biases [64], inequality [65], and racism [66].

Note that while such biases are not specifically situated in IoT systems, and there are IoT systems that do not interact with humans directly, such issues are relevant, as there are also IoT devices with internet applications that often employ face-recognition algorithms and voice-recognition algorithms (e.g., Google Home and Amazon Echo) and aim to present users with a summary of recent news and product recommendations.

Algorithmic bias can arise in autonomous systems [67], and could arise in IoT devices as they become increasingly autonomous. An IoT device that behaves using a machine-learning algorithm, if trained on bias data, could yield bias behaviour. With the increasing data-driven nature of IoT devices, a number of possible opportunities for discrimination can arise as noted in [68]—the examples given include an IoT gaming console and neighbourhood advisor that advises avoiding certain areas. Additionally, such an algorithmic bias can occur in machine-learning algorithms used for autonomous vehicles, where large volumes of data over time frames of minutes to days are analysed.

Even without using machine learning, it is not to difficult to think of devices that can exhibit biased behaviour—consider a sensor that is biased in the information it captures, intentionally or unintentionally, or a robot that greets certain types of people. Such a robot might be programmed to randomly choose who it greets, but it may happen to appear to only greet certain individuals and so be perceived as biased.

2.5.1. Racist Algorithms

While the algorithms or their developers might not be intentionally racist, as machine learning thrives on the data that they are trained on, bias can be introduced, even unintentionally. Hence, an algorithm may not be built to be intentionally racist but a failure of a face-recognition algorithm on those with a darker skin colour (for example, see https://www.ted.com/talks/joy_buolamwini_how_i_m_fighting_bias_in_algorithms?language=en, accessed on 1 July 2021) could raise concerns and cause a category of people to feel discriminated against. A device that has been trained to work in a certain context might not work as expected in a different context—a type of transfer bias. A simple example of a smart door trained to open based on recognizing fair-skinned persons might not operate for dark-skinned persons.

How IoT devices interface with humans could be biased by design, even if not intentionally so, but simply due to inadequate consideration.

As reported in the Technology Review on bias in natural language processing systems, (https://www.technologyreview.com/s/608619/ai-programs-are-learning-to-exclude-some-african-american-voices/, accessed on 1 July 2021), due to the use of machine learning to learn how to recognise speech, there are issues with minority population groups due to lack of training examples for the machine-learning algorithms. *"If there aren't enough examples of a particular accent or vernacular, then these systems may simply fail to understand you"*. The original intention and motive of developers could be considered when judging algorithmic bias, and care is needed to determine if and when bias does arise, even if not originally intended, especially with machine learning on data.

2.5.2. Other Algorithmic Bias

We have looked at how algorithms might appear to be racially biased in its inference but there could be other biases in general, for example, in politics, where the algorithm tends to favour a given political party more than others, or in business, where a particular

brand of goods is favoured over others. Suppose an algorithm used for recommending news articles or products for you does so in a systematically biased manner—it could then have an influence in your voting or buying behaviours. An algorithm that provides possibly biased recommendations or news is an issue that has put Facebook in the news, when it was said to "*deliberately suppressing conservative news from surfacing in its Trending Topics*" (http s://www.wired.com/2016/05/course-facebook-biased-thats-tech-works-today/, accessed on 1 July 2021). To provide greater transparency, Facebook has begun to describe how it recommends and filters news in the Trending Topics section (http://fortune.com/2016/0 5/12/facebook-and-the-news/, accessed on 1 July 2021), perhaps to be more open with the public. Other concerns are on how Twitter provides algorithmically filtered news feeds to users (http://fortune.com/2016/02/08/twitter-algorithm/, accessed on 1 July 2021).

What if the agenda is a "good" one, e.g., algorithms being informed by a utilitarian mandate? However, this raises the ethical question of whether software should be programmed to always benefit as many people as possible, even at the cost of a few—considering a hypothetical "smart" water-rationing system in homes, where water is conserved for all at the sacrifice of some urgent uses. In addition, taking a broader sustainability view, IoT systems can help cities move towards smarter, more energy-efficient homes, smarter waste management and smarter energy grids, helping to achieve sustainable development goals (for example, see https://deepblue.lib.umich.edu/bitstream/handle/2027.4 2/136581/Zhang_TheApplicationOfTheInternetOfThingsToEnhanceUrbanSustainability.p df?sequence=1&isAllowed=y , accessed on 1 July 2021 and http://www3.weforum.org/d ocs/IoTGuidelinesforSustainability.pdf, accessed on 1 July 2021). Another example is IoT-based infrastructure monitoring which helps to reduce urban flooding (https://www.wefo rum.org/agenda/2018/01/effect-technology-sustainability-sdgs-internet-things-iot/, accessed on 1 July 2021). However, how, in general, automated IoT systems should balance priorities within an overall sustainability agenda without bias towards or against any community groups, can be a consideration from the system design phase.

Moreover, there could be an issue with human values and bias being essentially incorporated into algorithms or into the design of IoT devices. The so-called *value-laden algorithms* by [69] can be defined as follows: "*An algorithm comprises an essential value-judgment if and only if, everything else being equal, software designers who accept different value-judgments would have a rational reason to design the algorithm differently (or choose different algorithms for solving the same problem).*" An example discussed is medical image algorithms. It is noted that "*medical image algorithms should be designed such that they are more likely to produce false positive rather than false negative results*". However, the increased number of false positives might lead to too many unnecessary operations. In addition, this could cause alarming results for many who are then suspected or diagnosed with diseases that they do not have, due to the medical image algorithms conservatively highlighting what is possibly not there. Hence, due to the need to be conservative and to avoid missing a diagnosis, a developer of the algorithm could have made it more pessimistic so that nothing is overlooked. Or, consider an internet camera that detects intruders which gives too many false positives in trying to be 'overly protective'.

To be fair, algorithmic bias can arise due to the developers own values, due to the data used in training algorithms, or simply due to cases not considered during design, and perhaps not due to malicious or intentionally biased agendas. However, an issue is how to distinguish between intentional (malicious) and unintentional algorithmic bias.

2.6. Issues with Cooperative IoT

When IoT devices cooperate, a number of issues arise. For example, with autonomous vehicles, it is not only vehicle-to-vehicle cooperation, as an autonomous vehicle could share roads with pedestrians, cyclists, and other human-driven vehicles, and would need to reason about social situations, perform social signalling with people via messaging or physical signs, and work within rules and norms for the road, which could prove to be a

difficult problem (https://spectrum.ieee.org/transportation/self-driving/the-big-proble m-with-selfdriving-cars-is-people, accessed on 1 July 2021).

Protection from false messages and groups of vehicles that cooperate maliciously, are also concerns, looking forward. How will a vehicle know if a message to make way is authentic? What if vehicles take turns to dominate parking spaces or gang-up to misinform non-gang vehicles about where available parking might be? Or what if vehicles of a given car manufacturer collude to the disadvantage other brands of cars?

A similar issue arises with other IoT devices, which must discern the truthfulness of messages they receive and, when cooperating and exchanging data, need to follow policies for data exchanges. Denial-of-Service attacks, where a device receives too many spurious messages, must be guarded against, and IoT devices should not spam other devices. The issues of trust with a large number of inter-connected devices has been explored with a proposed trust model, in [9].

With cooperation, considerations of what data should be shared and how data are shared among cooperative IoT devices will be important. For example, if a group of vehicles share routing information in order to cooperate on routing to reduce congestion, as in [70], there is a need to ensure that such information is not stored or used beyond their original purpose.

2.7. User Choice and Freedom

Apart from the transparency of operations, systems allowing adequate user choice is also important—freedom of action is an important property of systems that are respectful of the autonomy of users, or at least a user's direction is based on the user's own "*active assessment of factual information*" [71] (http://www.ethics.org.au/on-ethics/blog/october-2 016/ethics-explainer-autonomy, accessed on 1 July 2021). For example, a device can be programmed to collect data and manage data automatically (e.g., once a photo is taken by a device, it can be automatically shared with a number of parties and stored), but people would like to be informed about what data are collected and how data are used. Informing might not be adequate—a system could automatically inform a user that all photos on a smartphone will be copied to the cloud and will be categorised in a default manner on the phone, but the user might want control over what categories to use and control which photos should be copied to the cloud.

Another example is a smartphone that is programmed to only show the user certain Wi-Fi networks, restricting user choice, or a smartphone that filters out certain recommendations from applications, which can happen without the user's knowledge. In general, people would like to maintain choices and freedoms in the presence of automation—this is also discussed in the context of automated vehicles later.

In relation to location-based services, or more generally, context-aware mobile services, control and trust are concerns [72]. Someone might willingly give away location or contextual information in order to use particular services (an outcome of a privacy–utility trade off), assuming that she/he trusts the service; the user still retains the choice of opting in or not, and opting out anytime during the use of the service. Tracking a child for safety can be viewed as somewhat intruding on his/her privacy but might be insisted on by the parent. As mentioned in [72], in general, a wide range of considerations is required to judge whether such context-aware services are ethical or not ethical, including rules and norms, benefits and harms, concerns of people, governing bodies, and cultural values.

2.8. Summary and Discussion

We discuss a range of ethical concerns and issues in this section. From the above discussion, for an IoT device to behave ethically, ideally—minimally—the device should do the following:

- Communicate securely.
- Manage data in a way that is privacy-respecting.
- Act in a way that is sensitive to and aware of personal spaces and privacy concerns of users.
- Be able to make effective justifiable moral decisions depending on an appropriate ethical framework and guidelines.
- Act (if it is able to do so) in a way that is ethical when performing important functions that affect people (e.g., robots in health and social applications).
- Be deployed in an appropriate setting and environment in relation to its capabilities and function.
- Be free of algorithmic bias as far as this is relevant and feasible.
- Cooperate with other devices appropriately.
- Behave in ways that preserve user choice and freedom, and which invite and maintain user trust.
- Maintain transparency of operations as appropriate to different stakeholders.

In the above, what is "appropriate" will often be context-specific as we discuss later in the paper.

3. Towards a Multi-Pronged Approach

How one can build IoT devices that behave ethically is still a current area of research. This section reviews a range of ideas that have been proposed and applied to ameliorate the situation, including how to program ethical behaviour in devices, algorithmic transparency for accountability, algorithmic social contracts and crowdsourcing ethics solutions, enveloping IoT systems, and devising a code of ethics for developers. Then, it is argued that, as each idea has its own merits and usefulness towards addressing ethical concerns, a multi-pronged approach can be useful.

3.1. Programming Ethical Behaviour

We review a range of techniques that have been explored for programming ethical behaviour: rule-based programming and learning, game-theoretic calculations, ethical settings and ethical-by-design.

3.1.1. Rule-Based Ethics

If we want robots or an algorithm to be ethical and transparent, perhaps one way is to encode the required behaviour explicitly in rules or to create algorithms to allow devices to calculate the ethical actions. Rules have an advantage in that they can be human understandable, and they can represent a range of the ethical behaviours required of the machine. Foundational ideas of ethics, such as consequentialism and deontology, often underpin the approaches taken. The general idea is that a device whose behaviour abides by these rules is then considered ethical.

The work in [73] describes a vision of robots and an example of coding rules of behaviour for robots. *EthEL* [74] is an example of a robot that provides reminders about medication. There are issues of when to notify the (human) carer/overseer when the patient does not take the medication. It would be good for the patient to be respected and have a degree of autonomy to delay or not take medication, but an issue arises when, if the medication is not taken, it leads to a life-threatening situation—the issue is when the robot should persist in reminding and informing the overseer, or when it should not, respecting the autonomy of the patient.

A machine-learning algorithm based on inductive logic was used to learn a general ethical rule about what to do based on particular training cases given to the algorithm: "*a health care robot should challenge a patient's decision— violating the patient's autonomy—whenever doing otherwise would fail to prevent harm or severely violate the duty of promoting patient welfare*". In 2008, this was believed to be the first robot governed by an ethical principle, where the principle was learned.

The work in [75] proposes a framework for building ethical agents and the use of norms to characterise interactions among autonomous agents, with three types of norms, namely commitments, authorizations and prohibitions. Such norms can be used by agents needing to behave ethically. Such multiagent modelling maps well to decentralized IoT systems, allowing the placing of decentralised intelligence and algorithmic behaviour within the IoT [16].

Ethical questions can be much more complex, in general—it would be hard to encode in rules every conceivable situation where a robot should persist with its reminders, even when the patient rejects it. It remains an open research area as to what extent such rules can be coded up by a programmer, or learnt (via some machine learning algorithm) for machines in a diverse range of situations in other applications.

Another example is the work of [76] which outlines programming ethical behaviour for autonomous vehicles by mapping ethical considerations into costs and constraints used in automated control algorithms. Deontological rules are viewed as constraints in an optimal control problem of minimising costs, for example, in the case of deciding actions to minimise damages in an incident.

From the above examples, the general overarching rule is that saving human life takes priority over conforming to traffic laws and following a person's (perhaps under-informed) decision. However, it is generally difficult to ensure that a vehicle abides by these rules and for automated vehicles to assess situations accurately to know which rule applies. In addition, its software would need to be tested to follow such principles, or testing be done by a certification authority though requiring tremendous resources.

In [77], a formal model of safe and scalable self-driving cars is presented, where a set of behavioural rules are specified which can be followed by cars to help ensure safety. A rule-based approach can work for specific IoT applications, where the rules are identifiable and can be easily encoded.

However, in general, a difficulty is how one can comprehensively determine what the rules are for specific applications, apart from expert opinion. This raises the question of who decides what is ethical and what is not, and whether users could trust the developers who engineered the IoT systems on what is ethical behaviour. Apart from experts encoding rules, an alternative approach proposed by MIT researchers is to crowdsource human perspectives on moral decisions as experimented with by the Moral Machine for autonomous vehicles, with interesting results, including cross-cultural ethical variation [78] (http://moralmachine.mit.edu/, accessed on 1 July 2021).

System architectures for building machines capable of moral reasoning remains a research area [79,80] (also https://www.nature.com/news/machine-ethics-the-robot-s-dilemma-1.17881#auth-1, accessed on 1 July 2021). Recent work has proposed rule-based models to encode ethical behaviour that can be computationally verified [81], and in contrast to verification approaches, an internal simulation of actions by a machine in order to predict consequences is proposed in [82].

3.1.2. Game-Theoretic Calculation of Ethics

Game-theoretic approaches have also been proposed for autonomous vehicles to calculate ethical decisions, e.g., using Rawlsian principles in contrast to utilitarian approaches [83]. The idea is to determine, given the behaviour of the party, the best outcome. A difficulty is deciding whether Rawlsian or utilitarian calculation should be employed, or even other schemes—the Rawlsian approach aims to maximise utility for the worst case (maximin approach) while the utilitarian approach aims to maximise utility. It is also difficult to assign appropriate numerical values for the utilities of actions (e.g., why would hitting a pedestrian have a value of -1 while injuring a pedestrian is given a value of -0.5?).

3.1.3. Ethics Settings

Another category of work focuses on getting user input in 'programming' the ethical behaviour of devices, in particular, for autonomous cars. The notion of *ethics settings* or the 'ethical knob' was proposed by [84], to allow passengers of autonomous vehicles to make choices about ethical dilemmas, rather than have the reasoning hard-coded by manufacturers. For a vehicle needing to prioritise between the safety of the passengers and of pedestrians in road situations, there are three modes, namely *altruistic*, *egoistic* and *impartial*, corresponding to the preferences for the safety of the pedestrian, the safety of the passenger and the safety of both, and the passenger can choose a mode somewhere in between, among the three. The idea of the ethics settings is advocated by [85], which also answers the question of what settings people should use—each person choosing the selfish ethics settings might make society worst off, overall, while everyone, if this can be mandated, choosing the settings that minimise harm, even if altruistic, would make society better off.

3.1.4. Ethical by Design

In [86], an approach is to allow designers of IoT systems to configure via a set of available policy templates, which reduces the complexity of the software engineering of IoT systems, where multiple policies are relevant, e.g., a policy on storage of data, a policy on how data can be shared or a policy on ethical actions. A set of policies can be chosen by the user or the developer (in view of the user) that is tailored to the user's capabilities and context. A framework for dynamic IoT policy management is given in [87].

While this review does not focus on challenges of IoT data privacy specifically, the review in [88] noted that addressing IoT data privacy challenge involves designing and building in data access control and sharing mechanisms into IoT devices, e.g., building in authentication and key establishment mechanisms in IoT devices, computing on the edge to address privacy concerns, mechanisms to mask shared personal data, tools to support user access rules for data handling, and tools for digital forgetting and data summarization. In summary, one can reduce user privacy leakage and risks of IoT devices mishandling IoT data via a combination of these mechanisms.

When such mechanisms are known and existing, and as more such mechanisms are developed, then, according to [86], ethically designed IoT products (including devices and applications) are those "designed and deployed to empower users in controlling and protecting their personal data and any other information." The idea of the 'ethical knob' also seeks to put more control into the hands of users, beyond data handling. Hence, to program in ethical behaviour is not only programming IoT devices that take action based on ethical considerations, but also providing users appropriate control over device behaviour, even as it has delegated authority to act autonomously.

3.2. Enveloping IoT Systems

The concept of 'enveloping' was first introduced in [89] in regard to providing boundaries within which today's AI systems can work effectively. A distinction is made between complexity of a task, in relation to how much computational resources it requires, and the difficulty of the task, relating to the physical manipulation skills it requires, e.g., the gross or fine motor skills (robotic or human) required to perform tasks, such as dish washing with hands, painting with a brush, tying shoelaces, typing, using a tool, running up the stairs, playing an instrument, or helping the somewhat disabled walk or get up. Examples, taken from the paper of *envelopes*, for devices include, for industrial robotics, "the three-dimensional space that defines the boundaries within which a robot can work successfully is defined as the robot's envelope", the waterproof box of a dishwasher, and Amazon's robotic shelves and warehouse for its warehouse robots. It is noted that "driverless cars will become a commodity the day we can successfully envelop the environment around them." A computer chess program can be very successful within the constraints of the rules of chess. Indeed, the idea of dedicated lanes or areas for automated vehicles can be viewed

as a type of envelope for such vehicles. Hence, enveloping is a powerful idea for successful AI systems.

While it might not always be possible to envelop IoT systems, consider a generalized view of enveloping that is not just physical, but also cyber–physical, comprising the situation spaces (physical boundaries and cyber boundaries) in which a device functions. Such enveloping can help in addressing ethical issues by reducing the complexity of the environment in which such IoT devices or robots operate, reducing the chance for unintended situations, allowing comprehensive rules to be devised in a more constrained operating environment, helping to manage human expectations (e.g., humans generally get out of the way of trains, trams and vehicles on the road), and enabling clear definition of the context of operation, e.g., algorithmic bias is then not unexpected if the context of the development of the algorithm is known, such as the training data set used, and the internet environment or 'cyber-envelope' in which the device operates, including where data are stored and shared, is explicitly co-defined by IoT device manufacturers and users. Additionally, as another example, a pill-taking reminder system works within its known envelope so that unexpected behaviours when working beyond its envelope can be expected. However, enveloping can prove restrictive in the IoT, and successful enveloping to help deal with ethical IoT issues has yet to be proven.

3.3. White-Box Algorithms

As noted earlier, algorithms might be used to make decisions that affect people in a significant way, e.g., criminal cases and whether someone should be released from prison, to whether someone is diagnosed with a particular disease. In addition, certain groups of people may feel that it is unfair if an algorithm does not work as well for them as it does for someone else on account of his/her skin colour or accent.

How can one deal with algorithmic bias? Two areas of research to address this problem are noted: algorithmic transparency and detecting algorithmic bias.

3.3.1. Transparency

There are at least two aspects of transparency for IoT devices: the data traffic going into and out of such devices, and the inner workings of such devices.

For example, the TLS-RaR approach [90] allows device owners (or consumer watchdogs and researchers) to audit the traffic of their IoT devices. Affordable in-home devices called *auditors* can be added to the local network to observe network data for IoT devices, using Transport Layer Security (TLS) connections. However, some devices might use steganography to hide data in traffic, or users might still miss some data sent out by a malicious device.

Apart from monitoring the traffic of IoT devices, there are many who argue that algorithms that make important decisions should be a 'white box' rather than a 'black box' so that people can scrutinise and understand how the algorithms make decisions, and judge the algorithms that judge us. This is also a current emphasis for the explainability of AI idea (https://en.wikipedia.org/wiki/Explainable_Artificial_Intelligence, accessed on 1 July 2021) This can become an increasingly important feature for IoT devices that take action autonomously; users need to know why, for example, heating has been reduced in a smart home recently, forgetting that a target energy expenditure was set earlier in the month.

For widely used systems and devices where many people could be affected, transparency enables social accountability. IoT devices in public spaces, deployed by the town council, should work according to public expectations. IoT public street lights that systematically only light up certain segments of a road for particular shops and not for other shops can be seen to be biased—or at least, must be in error and be subject to scrutiny.

Consider IoT devices whose purpose is to provide information to people, or devices that filter and provide information to other devices; transparency in such devices enable people to understand (at least in part) why certain information is shown to them, or understand their behaviour. For example, Facebook has been rather open in how its newsfeed algorithm works (https://blog.bufferapp.com/facebook-news-feed-algorithm, accessed on 1 July 2021). It can be seen that by being open about how the algorithm works, Facebook provides, to an extent, a form of social accountability.

Another way an algorithm could 'expose' its workings is to output logs of its behaviour over time. For example, in the case of autonomous vehicles, if an accident happens between a human-driven car and an autonomous vehicle, one should be able to inspect logs to trace back to what happened and decide if the company should be held accountable or the human driver. This is similar to flight recorders in commercial airplanes. As another example of auditing, the Ditio [91] system is an approach for auditing sensor activities, logging activities that can be later inspected by an auditor and checking for compliance with policies. An example is given where the camera on a Nexus 5 smartphone is audited to ensure that it has not been used for recording during a meeting.

However, there are concerns with logging and white box views of algorithms. For example, intellectual property might be a concern when the workings of an algorithm is transparent, or when data used in training a machine-learning algorithm are exposed. Care must be taken in how algorithms are made transparent. Another issue is that, often, with neural network learning algorithms, the actual rules learnt for classification and decision making are also not explicitly represented (and are simply encoded in the parameters of the neural network). Additionally, what types of data or behaviour should be logged and how they can be managed remain open issues and are application-dependent.

The white-box algorithm approach can be employed to expose algorithmic bias when present or to allow human judgement on algorithmic decisions, but the workings of a complex algorithm are not easily legible or understandable in every situation.

Algorithms and systems may need to be *transparent by design*—a software engineering challenge. The paper on *Algorithmic Accountability* by the World Wide Web Foundation (http://webfoundation.org/docs/2017/07/Algorithms_Report_WF.pdf, accessed on 1 July 2021) calls for the explainability and auditability of software systems, particularly those based on machine learning, to encourage algorithmic accountability and fairness. The Association of Computing Machinery (ACM), a major computing association based in the U.S.A., calls for *algorithmic transparency and accountability* in a statement (https://www.acm.org/binaries/content/assets/public-policy/2017_usacm_statement_algorithms.pdf, accessed on 1 July 2021).

Getting algorithms or systems to explain their own actions and audit their own execution have become current research areas, as suggested by this workshop on Data and Algorithmic Transparency (http://datworkshop.org/#tab_program, accessed on 1 July 2021).

Additionally, using open-source software has been argued to be an approach to achieve transparency, e.g, of AI algorithms (https://www.linuxjournal.com/content/what-does-ethical-ai-mean-open-source, accessed on 1 July 2021). However, commercial interests might hinder free and open software.

In summary, for transparency and accountability, as noted in [92], IoT systems, from the technical point of view, can provide *control*—allowing users to control what happens, and *auditing*—enabling what happened or why something is happening to be recorded. IoT systems also need to allow users to understand (and perhaps control) what data they collect and what they do with that data to allow users to understand (and perhaps configure) their motivations (http://iot.stanford.edu/retreat15/sitp15-transparency-no.pdf, accessed on 1 July 2021), and to see (and perhaps change), in a non-technical way, how they work.

3.3.2. Detecting Algorithmic Bias

People might stumble upon such bias when using some devices, but sometimes it may be a lot more subtle (e.g., considering a news feed where we may not realize or miss what we are not expecting). Researchers have looked at how to detect algorithmic bias using systematic testing based on statistical methods, such as with a technique called *transparent model distillation* [93], which we will not look at in-depth here.

3.4. Black-Box Validation of Algorithmic Behaviour

There could be situations where white-boxing algorithms is not possible due to commercial reasons, and generating explanations from certain (e.g., deep learning) algorithms is still a challenge. It is well articulated in [94] that

> *"the study of machine behaviour will often require experimental intervention to study human–machine interactions in real-world settings".*

Software testing is well studied and practised. Experimental evaluation of algorithmic behaviour to verify certain properties or capabilities might be employed, though testing device behaviour in all circumstances and environments can be challenging, especially if it connects to other devices and if there are flow-on consequences in the physical world, considering the complexity of a device. A notion of the Turing test was proposed for autonomous vehicles (for example, see https://news.itu.int/a-driving-license-for-autonomous-vehicles-towards-a-turing-test-for-ai-on-our-roads/, accessed on 1 July 2021).

Where the range of possible situations and interactions with the environment are complex, apart from real-world testing, simulation-based testing and validation can be an economical solution as noted in [95], as an example. Software updates are expected to occur with IoT devices; validation might then need to be redone, changes localised to particular modules, and the impact of changes on other modules assessed. The work in [96,97] noted that the testing of autonomous vehicles and autonomous systems requires such *cognitive testing* as it is called.

The design of the human–device interface is also a consideration if users are to exercise choice and freedom, as they need to understand the functions of the device and how to interact with it. The interface should not be too complex so that users lose comprehension, yet should make available adequate choices and options. This is indeed a challenging task for a complex device. For example, for automated vehicle human–machine interfaces (HMIs), using heuristic evaluation is one approach [98], where a set of criteria is used to judge the HMI.

Validation requires criteria to validate against; a safety standards approach for fully autonomous vehicles is proposed in [99]. Similar standards of algorithmic behaviour might be devised for other types of IoT devices, e.g., for delivery robots on walkways, or robots in aged care homes.

3.5. Algorithmic Social Contracts

Going beyond the simple white-box concept for algorithms, the work in [100] proposed a conceptual framework for regulating AI and algorithmic systems.

The idea is to create social contracts between stakeholders that can hold an algorithmic system (and its developers) accountable and allow voting and negotiation on trade-offs (e.g., should a feature that increases pedestrian safety in autonomous vehicles but decrease passenger safety be implemented? Or should a feature of a system that decreases data privacy but increases public safety be deployed?). The idea is: *"to build institutions and tools that put the society in-the-loop of algorithmic systems, and allows us to program, debug, and monitor the algorithmic social contract between humans and governance algorithms".*

What is proposed is for tools to be developed that can take technical aspects of algorithms and present them to the general public so that they can be engaged in influencing the effect and behaviour of the algorithms, effectively *crowdsourcing ethics*, an approach used elsewhere [101]. The general approach of combining machine-learned representations

and human perspectives is also called *lensing* (https://www.media.mit.edu/videos/2017-05-18-karthik-dinakar/, accessed on 1 July 2021).

Tim O'Reilly's 2016 book, *Beyond Transparency: Open Data and the Future of Civic Innovation* proposes the idea of algorithmic regulation, (http://beyondtransparency.org/chapters/part-5/open-data-and-algorithmic-regulation/, accessed on 1 July 2021) where algorithmic regulation is successful when there is: "(1) a deep understanding of the desired outcome, (2) real-time measurement to determine if that outcome is being achieved, (3) algorithms (i.e., a set of rules) that make adjustments based on new data, and (4) periodic, deeper analysis of whether the algorithms themselves are correct and performing as expected". The actual process to achieve the above is still an unresolved socio-technical challenge, in itself an area of research.

3.6. Code of Ethics and Guidelines for IoT Developers

Rather than building ethical behaviour into machines, ethical guidelines are also useful for the developers of the technology. There are codes of ethics for robotics engineers (https://web.wpi.edu/Pubs/E-project/Available/E-project-030410-172744/unrestricted/A_Code_of_Ethics_for_Robotics_Engineers.pdf, accessed on 1 July 2021), and more recently, the Asilomar Principles for AI research. These principles were developed in conjunction with the 2017 Asilomar conference and relate to ethics in AI R&D (https://futureoflife.org/ai-principles/, accessed on 1 July 2021). The principles cover safety, transparency, privacy, incorporating human values and maintaining human control. The notion of how to imbue algorithms and systems with human values is a recent research topic (http://www.valuesincomputing.org/, accessed on 1 July 2021). The above appears to provide a morally sound path for AI R&D and AI applications, and for IoT devices with AI capabilities. Code of ethics for IoT is also being discussed (see the interview at https://www.theatlantic.com/technology/archive/2017/05/internet-of-things-ethics/524802/, accessed on 1 July 2021, and EU discussions at http://ec.europa.eu/transparency/regexpert/index.cfm?do=groupDetail.groupDetailDoc&id=7607&no=4, accessed on 1 July 2021). An IoT design manifesto (https://www.tudelft.nl/io/onderzoek/research-labs/connected-everyday-lab/iot-manifesto/, accessed on 1 July 2021) presents a range of general design principles for IoT developers. The IoT Alliance Australia has provided security guidelines (https://www.iot.org.au/wp/wp-content/uploads/2016/12/IoTAA-Security-Guideline-V1.2.pdf, accessed on 1 July 2021).

The German Federal Minister of Transport and Digital Infrastructure appointed a national ethics committee for automated and connected driving, who presented a code of ethics for automated and connected driving [102]. The ethical guidelines highlight a range of principles, including " ...*Technological development obeys the principle of personal autonomy, which means that individuals enjoy freedom of action for which they themselves are responsible*", i.e., personal autonomy is a key principle for ethical technology. Autonomous cars can improve the mobility of the disabled, and so, has ethical benefits. Another guideline stresses that official licensing and monitoring are required for automated driving systems, which may be a direction required for robotic autonomous things in public, from drones to delivery robots. A controversially debated guideline involves "..*General programming to reduce the number of personal injuries may be justifiable*", even at the cost of harm to some others, which somewhat adopts a utilitarian view of ethics, which may not be agreeable to all. Another guideline on accountability is: " *that manufacturers or operators are obliged to continuously optimize their systems and also to observe systems they have already delivered and to improve them where this is technologically possible and reasonable*", which is applied to automated vehicles but suggests possible implications on ethical IoT, in general, raising the question of the responsibility for maintenance and continual upgrades by manufacturers or operators for IoT devices post-deployment. Privacy-by-design is a principle suggested for data protection for connected driving.

It remains to be seen how ethical principles can be software engineered into future systems and whether certification requirements by law is possible, especially in relation to data handling by IoT devices (https://www.researchgate.net/publication/322628457_The _Legal_Challenges_of_Internet_of_Things, accessed on 1 July 2021).

The CTIA Cybersecurity Certification Test Plan for IoT devices (https://api.ctia.org/w p-content/uploads/2020/08/CTIA-IoT-Cybersecurity-Program-Management-Document -Ver-1.3.pdf, accessed on 1 July 2021) aims to define tests to be conducted on IoT devices for them to be certified in terms of three levels of security features built-in. Other standards and guidelines for IoT data privacy and device security are also being proposed and developed (https://www.schneier.com/blog/archives/2017/02/security_and_pr.html, accessed on 1 July 2021).

A comprehensive framework to help researchers identify and take into account challenges in ethics and law when researching IoT technologies, or more generally, heterogeneous systems, is given in [103]. Review of research projects by an ethics review board, consideration of national/regulatory guidelines and regulatory frameworks, and wider community engagement are among the suggested actions.

On a more philosophical note is the following question: what guidelines and strategies (or pro-social thinking) for the addition of each new device to the Internet of Things can encourage favourable evolution of the Internet of Things, even as it is being built? This is a generally challenging issue, especially in a competitive world, but the mechanisms of reciprocity, spatial selection, multilevel selection and kin selection are known to encourage cooperation [104]. Prosocial preferences sometimes do not explain human cooperation [105], and the question of how favourable human cooperation can happen continues to be explored, even from the viewpoint of models from statistical physics [106].

The work in [107] argues for embedding ethics into the development of robotics in healthcare via the concept of Responsible Research and Innovation (RRI). The RRI provides a toolkit of questions that helps to identify, understand and address ethical questions when applied to this domain.

3.7. Summary and Discussion

In summary, we make the following observations:

- *A multi-pronged approach*; Table 1 summarises the above discussion detailing the ideas and their main methods with their key advantages and technical challenges. It can be seen that each idea has advantages and challenges; they could complement each other so that combinations of ideas could be a way forward. Combining processes and artefact strategies would mean taking into account ethical guidelines and practices in the development of IoT devices and, where applicable, also building functionality into the device which allows the device to behave in an ethical manner (according to an agreed criteria) during operation. Devices can be built to work within the constraints of their enveloping environment, with user-informed limitations and clear expectations in terms of applicability, configurability, and behaviour. Developers could encode rules for ethical behaviour, but only after engagement and consultation with the community and stakeholders on what rules are relevant, based on a transparent and open process (e.g., consultative processes, technology trials, crowdsourcing viewpoints or online workshops). White- or gray-boxed devices could allow end-user intelligibility, consent and configurability, so that users retain a desired degree of control. Individual IoT devices should be secured against certain cyber-attacks, and the data they collect should be handled in a way that is intelligible and configurable by the user, according to best practice standards. When they take action, it should be in agreement with acceptable social norms, and auditable.

- *Context is key:* Many people are involved in an IoT ecosystem, including developers, IoT device retailers, IoT system administrators, IoT system maintainers, end-users, the local community and society at large. Society and communities can be affected by the deployment of such IoT systems in public, e.g., autonomous vehicles and robots in public; therefore, the broader context of deployment needs to be considered. Moreover, what is considered ethical behaviour might depend on the context of operation and the application; a device's actions might be considered ethical in one context but unethical in another, as also noted in [72] with regards to the use of location-based services. Broader contexts of operations include local culture, norms, or application domain (e.g., IoT in health, transport, or finance would have different rules for ethical behaviour); hence, it would require multiple levels of norms and ethical rules to guide the design and development of IoT devices and ecosystems. A basic ethical standard could apply (e.g., basic security built into devices, basic user-definable data handling options, and basic action tracking), with additional configurable options for context-specific ethical behaviour being added.
- *Ethical considerations with autonomy:* Guidelines for developers and consideration of what is built into an artefact to achieve ethical algorithmic behaviour could incorporate features that take into account, at least, the following:
  - Security of data and physical security as impacted by device actions.
  - Privacy of user data and device actions that impinge on privacy.
  - Consequences of over-reliance or human attachment to IoT devices.
  - Algorithmic bias and design bias, and fairness of device actions.
  - The possible need to engage not just end users, but anyone affected by the IoT deployment, e.g., via crowdsourcing viewpoints pre-development, and obtaining feedback from users and society at large, post-deployment.
  - User choice and freedom retained, including allowing user adjustments to ethical behaviour (e.g., opt in and out, adequate range of options, and designing devices with ethical settings).
  - End-user experience, including user intelligibility, scrutability and explainability, when needed, usability not just for certain groups of people, user control over data management and device behaviour, and appropriate manual overrides (https://cacm.acm.org/blogs/blog-cacm/238899-the-autocracy-of-autonomous-systems/fulltext, accessed on 1 July 2021).
  - Accountability for device actions, including legal and moral responsibilities, and support for traceability of actions.
  - Implications and possible unintended effects of cooperation among devices, e.g., where physical actions from multiple devices could mutually interfere, and the extent of data sharing during communications.
  - Deployment for long-term use (if applicable) and updatability, arising from security updates, improvements from feedback, adapting to changing human needs, policy changes, and
  - ethical consequences of autonomous action in IoT deployments (from physical movements to driving in certain ways).

Table 1. Summary of ideas to achieve ethical algorithmic behaviour in the IoT with key advantages and challenges.

	Idea	Methods	Key Advantages	Key Challenges	Selected Related Work
Build Behaviour into Artefact & Validate	Designing and Programming Ethical Behaviour	rule-based, game-theoretic calculations, ethics settings, ethical design templates	algorithmic or declarative representation of ethical behaviour, user control explicitly considered in artifact design	difficult for a set of rules to be complete, data used in development (e.g., to train Machine Learning models used in IoT devices) might be inadequate, hard for situations to be quantified, raises the question of who decides what is ethical	[16,73–75,77,83–87]
	Enveloping	setting physical/cyber-physical boundaries of operation	reduces complexity of operating environments, sets expectations in behaviour and contexts of trustworthy operation	may be hard to create suitable envelopes that do not hinder functioning of IoT systems	[89] (though originally proposed to achieve better AI systems)
	White-box Algorithms	improve transparency, detect algorithmic bias	greater traceability and accountability (*possibly allow engagement with non-developers*)	transparency does not equate understandability, scrutability does not equate user control,	[90–93]
	Black-box Validation	cognitive testing, simulation heuristic evaluation	where white-boxing is difficult, basis for certification	difficult to consider all cases and situations	[95,97,98]
	Algorithmic Social Contracts	crowdsourcing ethics, processes for algorithmic regulation	wider engagement (*possibly with non-developers*)	complex, may be hard to create suitable efficient processes or to obtain adequate participation	[100,101]
Guide Developers	Code of Ethics and Guidelines for IoT Developers	formal guidelines, regulations, community best practice for developers	highlights ethical considerations in development	application- or domain-specific considerations required	German ethics code for automated and connected driving [102], IoT data privacy guidelines and regulations [37], (also, Code of ethics for robotics engineers, Asilomar Principles, IoT design manifesto, IoT Alliance Australia Security Guideline, design-for-responsibility), RRI [107]

4. Conclusions and Future Work

This paper has reviewed a range of ethical concerns with IoT, including concerns that arise when IoT technology is combined with robotics and AI technology (including machine learning) and autonomous vehicles. The concerns include informational security, data privacy, moral dilemmas, roboethics, algorithmic bias when algorithms are used for decision making and control of IoT devices, as well as risks in cooperative IoT.

The paper also reviewed approaches that have been proposed to address these ethical concerns, including the following:

- Programming approaches to add ethical behaviour to devices, including adding moral reasoning capabilities to machines, and configuring devices with user ethics preferences.
- Detection and prevention of algorithmic bias, via accountability models and transparency.
- Behaviour-based validation techniques.
- The notion of algorithmic social contracts, and crowdsourcing solutions to ethical issues.
- The idea of enveloping systems, and
- developing guidelines and proposals for regulations, and codes of ethics, to encourage ethical developers and ethical development of IoT devices, and requiring security and privacy measures in such devices. Suitable data privacy laws in the IoT context, secure-by-design, privacy-by-design, ethical-by-design and design-for-responsibility principles are also needed.

A multi-pronged approach could be explored to achieve ethical IoT behaviour in a specific context. More research is required to explore combined approaches, and to create a framework of multiple levels of ethical rules and guidelines that could cater to the context-specific nature of what constitutes ethical behaviour. Indeed, by 'context-specific', it is meant taking into account a wide range of considerations in relation to the situation and operation of the device, e.g., the location, who is around, who the user is, the local regulations, the relevant societal and cultural values, the application, the function of the device, the capabilities of the device and the current physical environment.

This paper did not consider in detail the legislation and the laws involving robots and AI and the approaches of which could be considered for intelligent IoT systems, which are addressed in depth elsewhere [108]. The notion of IoT policing was not discussed in the sense of the run-time monitoring of devices to detect misbehaving devices, perhaps with the use of sentinel devices, as well as policy enforcement and penalties imposed on anti-social IoT devices (e.g., game-theoretic grim-trigger type strategies, and other types of sanctions for autonomous systems [109]). Social equity and social inequality are two concerns of the social ethics of the Internet of Things which are discussed elsewhere [110] but not detailed here. Sustainability of IoT deployments [111] and the use of IoT for sustainability [112] were not extensively discussed here, which have socio-ethical implications.

The challenge of building ethical things in the IoT that act autonomously yet ethically will also benefit from on-going research in building ethics into AI decision-making as reviewed in [20], which includes individual ethical decision frameworks, collective ethical decision frameworks, ethics in human–AI interactions and systems to explore ethical dilemmas. Future developments in topics covered in detail elsewhere (e.g., [113]) such as algorithmic fairness, differential privacy and game theory, will also help enable ethical algorithmic behaviour in IoT devices.

Outstanding socio-technical challenges remain if IoT devices are to behave ethically and be used ethically for IoT developers and IoT users. Ethical considerations would need to be factored into future IoT software and hardware development processes, according to upcoming certification practices, ethics policies, and regulatory frameworks, which are still to be developed. Particular domains or contexts require domain-specific guidelines and ethical considerations.

While we addressed mainly ethical behaviour for IoT device operations and the algorithms therein, there are ethical issues concerning the post-deployment and maintenance of IoT devices, where retailers or manufacturers could take responsibility.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Hermsen, S.; Frost, J.H.; Robinson, E.; Higgs, S.; Mars, M.; Hermans, R.C.J. Evaluation of a Smart Fork to Decelerate Eating Rate. *J. Acad. Nutr. Diet.* **2016**, *116*, 1066–1067. [CrossRef]
2. Kadomura, A.; Li, C.Y.; Chen, Y.C.; Tsukada, K.; Siio, I.; Chu, H.H. Sensing Fork: Eating Behavior Detection Utensil and Mobile Persuasive Game. In *Proceedings of the CHI'13 Extended Abstracts on Human Factors in Computing Systems*; ACM: New York, NY, USA, 2013; pp. 1551–1556. [CrossRef]
3. Wu, Q.; Ding, G.; Xu, Y.; Feng, S.; Du, Z.; Wang, J.; Long, K. Cognitive Internet of Things: A New Paradigm Beyond Connection. *IEEE Internet Things J.* **2014**, *1*, 129–143. [CrossRef]
4. Minerva, R.; Biru, A.; Rotondi, D. IEEE Internet Initiative. Available online: https://internetinitiative.ieee.org/ (accessed on 1 July 2021). .
5. Pintus, A.; Carboni, D.; Piras, A. Paraimpu: A Platform for a Social Web of Things. In Proceedings of the 21st International Conference on World Wide Web, Lyon, France, 16–20 April 2012. Association for Computing Machinery: New York, NY, USA, 2012; pp. 401–404. [CrossRef]
6. Taivalsaari, A.; Mikkonen, T. A Roadmap to the Programmable World: Software Challenges in the IoT Era. *IEEE Softw.* **2017**, *34*, 72–80. [CrossRef]
7. Tripathy, B.K.; Dutta, D.; Tazivazvino, C., On the Research and Development of Social Internet of Things. In *Internet of Things (IoT) in 5G Mobile Technologies*; Mavromoustakis, C.X., Mastorakis, G., Batalla, J.M., Eds.; Springer International Publishing: Cham, Switzerland, 2016; pp. 153–173._8. [CrossRef]

8.   Pticek, M.; Podobnik, V.; Jezic, G. Beyond the Internet of Things: The Social Networking of Machines. *Int. J. Distrib. Sens. Netw.* **2016**, *12*, 8178417. [CrossRef]

9.   Lin, Z.; Dong, L. Clarifying Trust in Social Internet of Things. *CoRR* **2017**, CrossRef]

10.   Farris, I.; Girau, R.; Militano, L.; Nitti, M.; Atzori, L.; Iera, A.; Morabito, G. Social Virtual Objects in the Edge Cloud. *IEEE Cloud Comput.* **2015**, *2*, 20–28. [CrossRef]

11.   Loke, S. *Context-Aware Pervasive Systems*; Auerbach Publications: Boston, MA, USA, 2006.

12.   Cristea, V.; Dobre, C.; Pop, F. Context-Aware Environments for the Internet of Things. In *Internet of Things and Inter-Cooperative Computational Technologies for Collective Intelligence*; Bessis, N., Xhafa, F., Varvarigou, D., Hill, R., Li, M., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 25–49._2. [CrossRef]

13.   Berman, F.; Cerf, V.G. Social and Ethical Behavior in the Internet of Things. *Commun. ACM* **2017**, *60*, 6–7. [CrossRef]

14.   Atlam, H.F.; Wills, G.B. *IoT Security, Privacy, Safety and Ethics*; Springer International Publishing: Cham, Switzerland, 2020; pp. 123–149.

15.   Calvo, P. The ethics of Smart City (EoSC): Moral implications of hyperconnectivity, algorithmization and the datafication of urban digital society. *Ethics Inf. Technol.* **2020**, *22*, 141–149. [CrossRef]

16.   Singh, M.P.; Chopra, A.K. The Internet of Things and Multiagent Systems: Decentralized Intelligence in Distributed Computing. In Proceedings of the 2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS), Atlanta, GA, USA, 5–8 June 2017; pp. 1738–1747. [CrossRef]

17.   Lipson, H.; Kurman, M. *Driverless: Intelligent Cars and the Road Ahead*; MIT Press: Cambridge, MA, USA, 2016.

18.   Simoens, P.; Dragone, M.; Saffiotti, A. The Internet of Robotic Things: A review of concept, added value and applications. *Int. J. Adv. Robot. Syst.* **2018**, *15*, 1729881418759424. [CrossRef]

19.   Ray, P.P. Internet of Robotic Things: Concept, Technologies, and Challenges. *IEEE Access* **2016**, *4*, 9489–9500. [CrossRef]

20.   Yu, H.; Shen, Z.; Miao, C.; Leung, C.; Lesser, V.R.; Yang, Q. Building Ethics into Artificial Intelligence. In Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI-18, Stockholm, Sweden, 13–19 July 2018; pp. 5527–5533.

21.   Stahl, B.C.; Timmermans, J.; Mittelstadt, B.D. The Ethics of Computing: A Survey of the Computing-Oriented Literature. *ACM Comput. Surv.* **2016**, *48*, 55:1–55:38. [CrossRef]

22.   Allhoff, F.; Henschke, A. The Internet of Things: Foundational ethical issues. *Internet Things* **2018**, *1–2*, 55–66. [CrossRef]

23.   Karale, A. The Challenges of IoT Addressing Security, Ethics, Privacy, and Laws. *Internet Things* **2021**, *15*, 100420. [CrossRef]

24.   Sha, K.; Wei, W.; Yang, T.A.; Wang, Z.; Shi, W. On security challenges and open issues in Internet of Things. *Future Gener. Comput. Syst.* **2018**, *83*, 326–337. [CrossRef]

25.   Ge, M.; Hong, J.B.; Guttmann, W.; Kim, D.S. A framework for automating security analysis of the internet of things. *J. Netw. Comput. Appl.* **2017**, *83*, 12–27. [CrossRef]

26.   Sicari, S.; Rizzardi, A.; Grieco, L.; Coen-Porisini, A. Security, privacy and trust in Internet of Things: The road ahead. *Comput. Netw.* **2015**, *76*, 146–164. [CrossRef]

27.   Malina, L.; Hajny, J.; Fujdiak, R.; Hosek, J. On perspective of security and privacy-preserving solutions in the internet of things. *Comput. Netw.* **2016**, *102*, 83–95. [CrossRef]

28.   Atwady, Y.; Hammoudeh, M. A Survey on Authentication Techniques for the Internet of Things. In Proceedings of the International Conference on Future Networks and Distributed Systems, Cambridge, UK, 19–20 July 2017; ACM: New York, NY, USA, 2017. [CrossRef]

29.   Sfar, A.R.; Natalizio, E.; Challal, Y.; Chtourou, Z. A roadmap for security challenges in the Internet of Things. *Digit. Commun. Netw.* **2018**, *4*, 118–137. [CrossRef]

30.   Alaba, F.A.; Othman, M.; Hashem, I.A.T.; Alotaibi, F. Internet of Things security: A survey. *J. Netw. Comput. Appl.* **2017**, *88*, 10–28. [CrossRef]

31.   Trnka, M.; Cerny, T.; Stickney, N. Survey of Authentication and Authorization for the Internet of Things. *Secur. Commun. Netw.* **2018**, *2018*, 4351603. [CrossRef]

32.   Díaz, M.; Martín, C.; Rubio, B. State-of-the-art, challenges, and open issues in the integration of Internet of things and cloud computing. *J. Netw. Comput. Appl.* **2016**, *67*, 99–117. [CrossRef]

33.   Jayaraman, P.P.; Yang, X.; Yavari, A.; Georgakopoulos, D.; Yi, X. Privacy preserving Internet of Things: From privacy techniques to a blueprint architecture and efficient implementation. *Future Gener. Comput. Syst.* **2017**, *76*, 540–549. [CrossRef]

34.   Weinberg, B.D.; Milne, G.R.; Andonova, Y.G.; Hajjat, F.M. Internet of Things: Convenience vs. privacy and secrecy. *Bus. Horizons* **2015**, *58*, 615–624. [CrossRef]

35.   Caron, X.; Bosua, R.; Maynard, S.B.; Ahmad, A. The Internet of Things (IoT) and its impact on individual privacy: An Australian perspective. *Comput. Law Secur. Rev.* **2016**, *32*, 4–15. [CrossRef]

36.   Weber, R.H. Internet of things: Privacy issues revisited. *Comput. Law Secur. Rev.* **2015**, *31*, 618–627. [CrossRef]

37.   Vegh, L. A Survey of Privacy and Security Issues for the Internet of Things in the GDPR Era. In Proceedings of the 2018 International Conference on Communications (COMM), Bucharest, Romania, 14–16 June 2018; pp. 453–458. [CrossRef]

38.   Humayed, A.; Lin, J.; Li, F.; Luo, B. Cyber-Physical Systems Security—A Survey. *IEEE Internet Things J.* **2017**, *4*, 1802–1831. [CrossRef]

39. Dutta, S. Striking a Balance between Usability and Cyber-Security in IoT Devices. Master's Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2017.

40. Halperin, D.; Heydt-Benjamin, T.S.; Ransford, B.; Clark, S.S.; Defend, B.; Morgan, W.; Fu, K.; Kohno, T.; Maisel, W.H. Pacemakers and Implantable Cardiac Defibrillators: Software Radio Attacks and Zero-Power Defenses. In Proceedings of the 2008 IEEE Symposium on Security and Privacy (SP 2008), Oakland, CA, USA, 18–22 May 2008; pp. 129–142.

41. Mittelstadt, B. Ethics of the health-related internet of things: A narrative review. *Ethics Inf. Technol.* **2017**, *19*, 157–175. [CrossRef]

42. Chamberlain, A.; Crabtree, A.; Haddadi, H.; Mortier, R. Special theme on privacy and the Internet of things. *Pers. Ubiquitous Comput.* **2017**, *22*, 289–292. [CrossRef]

43. Popescul, D.; Georgescu, M. Internet of Things—Some Ethical Issues. *USV Ann. Econ. Public Adm.* **2013**, *13*, 210–216.

44. Ali, M.S.; Dolui, K.; Antonelli, F. IoT Data Privacy via Blockchains and IPFS. In Proceedings of the Seventh International Conference on the Internet of Things, Linz, Austria, 22–25 October 2017; ACM: New York, NY, USA, 2017; pp. 14:1–14:7. [CrossRef]

45. Fernandez-Carames, T.M.; Fraga-Lamas, P. A Review on the Use of Blockchain for the Internet of Things. *IEEE Access* **2018**, *6*, 32979–33001. [CrossRef]

46. Griggs, K.N.; Ossipova, O.; Kohlios, C.P.; Baccarini, A.N.; Howson, E.A.; Hayajneh, T. Healthcare Blockchain System Using Smart Contracts for Secure Automated Remote Patient Monitoring. *J. Med. Syst.* **2018**, *42*, 130. [CrossRef]

47. Reyna, A.; Martín, C.; Chen, J.; Soler, E.; Díaz, M. On blockchain and its integration with IoT. Challenges and opportunities. *Future Gener. Comput. Syst.* **2018**, *88*, 173–190. [CrossRef]

48. Yu, B.; Wright, J.; Nepal, S.; Zhu, L.; Liu, J.; Ranjan, R. IoTChain: Establishing Trust in the Internet of Things Ecosystem Using Blockchain. *IEEE Cloud Comput.* **2018**, *5*, 12–23. [CrossRef]

49. Fister, I.; Ljubič, K.; Suganthan, P.N.; Perc, M.; Fister, I. Computational intelligence in sports: Challenges and opportunities within a new research domain. *Appl. Math. Comput.* **2015**, *262*, 178–186. [CrossRef]

50. Lima, A.; Rocha, F.; Völp, M.; Esteves-Veríssimo, P. Towards Safe and Secure Autonomous and Cooperative Vehicle Ecosystems. In *Proceedings of the 2Nd ACM Workshop on Cyber-Physical Systems Security and Privacy*; ACM: New York, NY, USA, 2016; pp. 59–70. [CrossRef]

51. Nasser, A.M.; Ma, D.; Muralidharan, P. An Approach for Building Security Resilience in AUTOSAR Based Safety Critical Systems. *J. Cyber Secur. Mobil.* **2017**, *6*, 271–304. [CrossRef]

52. Nawrath, T.; Fischer, D.; Markscheffel, B. Privacy-sensitive data in connected cars. In Proceedings of the 2016 11th International Conference for Internet Technology and Secured Transactions (ICITST), Barcelona, Spain, 5–7 December 2016; pp. 392–393. [CrossRef]

53. Bonnefon, J.F.; Shariff, A.; Rahwan, I. The social dilemma of autonomous vehicles. *Science* **2016**, *352*, 1573–1576. [CrossRef]

54. Lin, P., Why Ethics Matters for Autonomous Cars. In *Autonomous Driving: Technical, Legal and Social Aspects*; Maurer, M., Gerdes, J.C., Lenz, B., Winner, H., Eds.; Springer: Berlin/Heidelberg, Germany, 2016; pp. 69–85._4. [CrossRef]

55. Sparrow, R.; Howard, M. When human beings are like drunk robots: Driverless vehicles, ethics, and the future of transport. *Transp. Res. Part C Emerg. Technol.* **2017**, *80*, 206–215. [CrossRef]

56. Bagloee, S.A.; Tavana, M.; Asadi, M.; Oliver, T. Autonomous vehicles: Challenges, opportunities, and future implications for transportation policies. *J. Mod. Transp.* **2016**, *24*, 284–303. [CrossRef]

57. Lin, P.; Abney, K.; Bekey, G.A. *Robot Ethics: The Ethical and Social Implications of Robotics*; The MIT Press: Cambridge, MA, USA, 2014.

58. Lin, P.; Abney, K.; Jenkins, R. *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence*; Oxford University Press: Oxford, UK, 2017. Available online: https://oxford.universitypressscholarship.com/view/10.1093/oso/9780190652951.001.0001/oso-9780190652951 (accessed on 1 July 2021).

59. Tzafestas, S.G. *Roboethics: A Navigating Overview*; Springer: Berlin/Heidelberg, Germany, 2016.

60. Gunkel, D.J. The other question: Can and should robots have rights? *Ethics Inf. Technol.* **2017**, *20*, 87–99. . [CrossRef]

61. Cai, X.; Ning, H.; Dhelim, S.; Zhou, R.; Zhang, T.; Xu, Y.; Wan, Y. Robot and its living space: A roadmap for robot development based on the view of living space. *Digit. Commun. Netw.* **2020**. Available online: https://www.sciencedirect.com/science/article/pii/S2352864820302881 (accessed on 1 July 2021).

62. Enemark, C. *Armed Drones and the Ethics of War: Military Virtue in a Post-Heroic Age*; Routledge. 2013. Available online: https://www.semanticscholar.org/paper/Armed-Drones-and-the-Ethics-of-War%3A-Military-virtue-Enemark/7126533ca42895dada35ac0106d1c3956ab32e8b (accessed on 1 July 2021).

63. Mittelstadt, B.; Allo, P.; Taddeo, M.; Wachter, S.; Floridi, L. The ethics of algorithms: Mapping the debate. *Big Data Soc.* **2016**, *3*, 2053951716679679. [CrossRef]

64. O'Neil, C. *Weapons of Math Destruction*; Crown Books. 2016. Available online: https://www.amazon.com/Weapons-Math-Destruction-Increases-Inequality/dp/0553418815 (accessed on 1 July 2021).

65. Eubanks, V. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*; St. Martin's Press. 2018. Available online: https://virginia-eubanks.com/books/ (accessed on 1 July 2021).

66. Noble, S.U. *Algorithms of Oppression: How Search Engines Reinforce Racism*; NYU Press. 2018. Available online: https://www.tandfonline.com/doi/abs/10.1080/01419870.2019.1635260?journalCode=rers20 (accessed on 1 July 2021).

67. Danks, D.; London, A.J. Algorithmic Bias in Autonomous Systems. In Proceedings of the 26th International Joint Conference on Artificial Intelligence, Melbourne, Australia, 19–25 August 2017; pp. 4691–4697.

68. Tschider, C.A. Regulating the IoT: Discrimination, Privacy, and Cybersecurity in the Artificial Intelligence Age. *Denver Univ. Law Rev.* **2018**. [CrossRef]

69. Kraemer, F.; van Overveld, K.; Peterson, M. Is there an ethics of algorithms? *Ethics Inf. Technol.* **2011**, *13*, 251–260. [CrossRef]

70. Desai, P.; Loke, S.W.; Desai, A.; Singh, J. CARAVAN: Congestion Avoidance and Route Allocation Using Virtual Agent Negotiation. *IEEE Trans. Intell. Transp. Syst.* **2013**, *14*, 1197–1207. [CrossRef]

71. May, T. The Concept of Autonomy. *Am. Philos. Q.* **1994**, *31*, 133–144.

72. Abbas, R.; Michael, K.; Michael, M. Using a Social-Ethical Framework to Evaluate Location-Based Services in an Internet of Things World. *Int. Rev. Inf. Ethics* **2014**, *22*, 42–73.

73. Anderson, M.; Anderson, S.L. Robot be good. *Sci. Am.* **2010**, *303*, 72–77. [CrossRef]

74. Anderson, M.; Anderson, S.L. ETHEL: Toward a Principled Ethical Eldercare System. In Proceedings of the AI in Eldercare: New Solutions to Old Problems, Papers from the 2008 AAAI Fall Symposium, Arlington, VA, USA, 7–9 November 2008; pp. 4–11.

75. Ajmeri, N.; Guo, H.; Murukannaiah, P.K.; Singh, M.P. Designing Ethical Personal Agents. *IEEE Internet Comput.* **2018**, *22*, 16–22. [CrossRef]

76. Gerdes, J.C.; Thornton, S.M. Implementable Ethics for Autonomous Vehicles. In *Autonomous Driving: Technical, Legal and Social Aspects*; Maurer, M, Gerdes, J.C., Lenz, B., Winner, H., Eds.; Springer: Berlin/Heidelberg, Germany, 2016; pp. 87–102._5. [CrossRef]

77. Shalev-Shwartz, S.; Shammah, S.; Shashua, A. On a Formal Model of Safe and Scalable Self-driving Cars. *CoRR* **2017**. Available online: https://export.arxiv.org/pdf/1708.06374 (accessed on 1 July 2021).

78. Awad, E.; Dsouza, S.; Kim, R.; Schulz, J.; Henrich, J.; Shariff, A.; Bonnefon, J.F.; Rahwan, I. The Moral Machine experiment. *Nature* **2018**, *563*, 59–64. [CrossRef] [PubMed]

79. Wallach, W.; Allen, C. *Moral Machines: Teaching Robots Right from Wrong*; Oxford University Press, Inc.: New York, NY, USA, 2010.

80. Anderson, M.; Anderson, S.L. *Machine Ethics*, 1st ed.; Cambridge University Press: New York, NY, USA, 2011.

81. Dennis, L.; Fisher, M.; Slavkovik, M.; Webster, M. Formal verification of ethical choices in autonomous systems. *Robot. Auton. Syst.* **2016**, *77*, 1–14. [CrossRef]

82. Vanderelst, D.; Winfield, A. An architecture for ethical robots inspired by the simulation theory of cognition. *Cogn. Syst. Res.* **2018**, *48*, 56–66. [CrossRef]

83. Leben, D. A Rawlsian Algorithm for Autonomous Vehicles. *Ethics Inf. Technol.* **2017**, *19*, 107–115. [CrossRef]

84. Contissa, G.; Lagioia, F.; Sartor, G. The Ethical Knob: Ethically-customisable automated vehicles and the law. *Artif. Intell. Law* **2017**, *25*, 365–378. [CrossRef]

85. Gogoll, J.; Müller, J.F. Autonomous Cars: In Favor of a Mandatory Ethics Setting. *Sci. Eng. Ethics* **2017**, *23*, 681–700. [CrossRef] [PubMed]

86. Baldini, G.; Botterman, M.; Neisse, R.; Tallacchini, M. Ethical Design in the Internet of Things. *Sci. Eng. Ethics* **2016**. [CrossRef] [PubMed]

87. Sicari, S.; Rizzardi, A.; Miorandi, D.; Coen-Porisini, A. Dynamic Policies in Internet of Things: Enforcement and Synchronization. *IEEE Internet Things J.* **2017**, *4*, 2228–2238. [CrossRef]

88. Seliem, M.; Elgazzar, K.; Khalil, K. Towards Privacy Preserving IoT Environments: A Survey. *Wirel. Commun. Mob. Comput.* **2018**, *2018*, 1032761. [CrossRef]

89. Floridi, L. What the Near Future of Artificial Intelligence Could Be. *Philos. Technol.* **2019**, *32*, 1–15. [CrossRef]

90. Wilson, J.; Wahby, R.S.; Corrigan-Gibbs, H.; Boneh, D.; Levis, P.; Winstein, K. Trust but Verify: Auditing the Secure Internet of Things. In *Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services*; ACM: New York, NY, USA, 2017; pp. 464–474. [CrossRef]

91. Mirzamohammadi, S.; Chen, J.A.; Sani, A.A.; Mehrotra, S.; Tsudik, G. Ditio: Trustworthy Auditing of Sensor Activities in Mobile & IoT Devices. In Proceedings of the 15th ACM Conference on Embedded Network Sensor Systems, Delft, The Netherlands, 6–8 November 2017; ACM: New York, NY, USA, 2017; pp. 28:1–28:14. [CrossRef]

92. Singh, J.; Millard, C.; Reed, C.; Cobbe, J.; Crowcroft, J. Accountability in the IoT: Systems, Law, and Ways Forward. *Computer* **2018**, *51*, 54–65. [CrossRef]

93. Tan, S.; Caruana, R.; Hooker, G.; Lou, Y. Detecting Bias in Black-Box Models Using Transparent Model Distillation. *arXiv* **2017**, arXiv:1710.06169.

94. Rahwan, I.; Cebrian, M.; Obradovich, N.; Bongard, J.; Bonnefon, J.F.; Breazeal, C.; Crandall, J.; Christakis, N.; Couzin, I.; Jackson, M.; et al. Machine behaviour. *Nature* **2019**, *568*, 477–486. [CrossRef]

95. Kufieta, K.; Ditze, M. A virtual environment for the development and validation of highly automated driving systems. In *17. Internationales Stuttgarter Symposium*; Bargende, M., Reuss, H.C., Wiedemann, J., Eds.; Springer Fachmedien Wiesbaden: Wiesbaden, Germany, 2017; pp. 1391–1401.

96. Ebert, C.; Weyrich, M. Validation of Automated and Autonomous Vehicles. *ATZelectronics Worldw.* **2019**, *14*, 26–31. [CrossRef]

97. Ebert, C.; Weyrich, M. Validation of Autonomous Systems. *IEEE Softw.* **2019**, *36*, 15–23. [CrossRef]

98. Naujoks, F.; Wiedemann, K.; Schömig, N.; Hergeth, S.; Keinath, A. Towards guidelines and verification methods for automated vehicle HMIs. *Transp. Res. Part F Traffic Psychol. Behav.* **2019**, *60*, 121–136. [CrossRef]

99. Koopman, P.; Ferrell, U.; Fratrik, F.; Wagner, M. A Safety Standard Approach for Fully Autonomous Vehicles. In *Computer Safety, Reliability, and Security*; Romanovsky, A., Troubitsyna, E., Gashi, I., Schoitsch, E., Bitsch, F., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 326–332.

100. Rahwan, I. Society-in-the-loop: Programming the algorithmic social contract. *Ethics Inf. Technol.* **2017**. [CrossRef]

101. Lieberman, H.; Dinakar, K.; Jones, B. Crowdsourced Ethics with Personalized Story Matching. In Proceedings of the CHI'13 Extended Abstracts on Human Factors in Computing Systems, Paris, France, 27 April–2 May 2013 ; ACM: New York, NY, USA, 2013; pp. 709–714. [CrossRef]

102. Luetge, C. The German Ethics Code for Automated and Connected Driving. *Philos. Technol.* **2017**, *30*, 547–558. [CrossRef]

103. Happa, J.; Nurse, J.R.C.; Goldsmith, M.; Creese, S.; Williams, R. An ethics framework for research into heterogeneous systems. In Proceedings of the Living in the Internet of Things: Cybersecurity of the IoT—2018, London, UK, 28–29 March 2018; pp. 1–8. [CrossRef]

104. Rand, D.G.; Nowak, M.A. Human cooperation. *Trends Cogn. Sci.* **2013**, *17*, 413–425. [CrossRef] [PubMed]

105. Burton-Chellew, M.N.; West, S.A. Prosocial preferences do not explain human cooperation in public-goods games. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 216–221. [CrossRef] [PubMed]

106. Perc, M.; Jordan, J.J.; Rand, D.G.; Wang, Z.; Boccaletti, S.; Szolnoki, A. Statistical physics of human cooperation. *Phys. Rep.* **2017**, *687*, 1–51. [CrossRef]

107. Stahl, B.C.; Coeckelbergh, M. Ethics of healthcare robotics: Towards responsible research and innovation. *Robot. Auton. Syst.* **2016**, *86*, 152–161. [CrossRef]

108. Pagallo, U. *The Laws of Robots—Crimes, Contracts, and Torts*; Law, Governance and Technology Series; Springer: Berlin/Heidelberg, Germany, 2013; Volume 10. [CrossRef]

109. Nardin, L.G.; Balke-Visser, T.; Ajmeri, N.; Kalia, A.K.; Sichman, J.S.; Singh, M.P. Classifying sanctions and designing a conceptual sanctioning process model for socio-technical systems. *Knowl. Eng. Rev.* **2016**, *31*, 142–166. [CrossRef]

110. Shahraki, A.; Haugen, ø. Social ethics in Internet of Things: An outline and review. In Proceedings of the 2018 IEEE Industrial Cyber-Physical Systems (ICPS) Conference, Saint Petersburg, Russia, 15–18 May 2018; pp. 509–516. [CrossRef]

111. Stead, M.; Coulton, P.; Lindley, J.; Coulton, C. *The Little Book of SUSTAINABILITY for the Internet of Things*. 2019. Available online: https://www.researchgate.net/publication/331114232_The_Little_Book_of_SUSTAINABILITY_for_the_Internet_of_T hings (accessed on 1 July 2021).

112. Bibri, S.E. The IoT for smart sustainable cities of the future: An analytical framework for sensor-based big data applications for environmental sustainability. *Sustain. Cities Soc.* **2018**, *38*, 230–253. [CrossRef]

113. Kearns, M.; Roth, A. *The Ethical Algorithm: The Science of Socially Aware Algorithm Design*; Oxford University Press: Oxford, UK, 2019.