Office for Product Safety & Standards

Study on the Impact of Artificial Intelligence on Product Safety

Final Report

December 2021

A report for the Office for Product Safety and Standards (OPSS) by the Centre for Strategy and Evaluation Services (CSES)

Acknowledgements

This independent research report was produced by CSES.

The views expressed in this report are those of the authors, not necessarily those of the Office for Product Safety and Standards or the Department for Business, Energy & Industrial Strategy (nor do they reflect Government policy).

We are grateful to all respondents for their time and insights (See Appendix D for a list of stakeholders consulted).

Contents

Executive Summary

The study on the impact of artificial intelligence on product safety was commissioned by the Office for Product Safety and Standards (OPSS), part of the Department for Business, Energy and Industrial Strategy (BEIS), and carried out by the Centre for Strategy and Evaluation Services (CSES) between January and June 2021.

Below we first provide an overview of the study objectives and methodology, before presenting a summary of the study's main findings and conclusions.

Study objectives and methodology

The objective of this study was to **examine the current and forecasted future impacts of artificial intelligence (AI) in consumer products, and what this means for product safety**. This breaks down into the following three specific objectives:

- **Objective 1:** Analyse the current and likely future applications of AI in the home, highlighting the advantages and disadvantages for consumers and product safety implications / risks.
- **Objective 2:** Assess whether the current product safety framework is sufficient for a new generation of products that incorporate AI.
- **Objective 3:** Examine what factors Regulators should consider when responding to these new challenges to ensure consumer safety and foster product innovation.

Considering the **study scope**, all manufactured consumer products subject to the General Product Safety Regulations (GPSR) 2005 and other relevant legislation for specific goods are covered, except for vehicles, pharmaceuticals and food. Cyber security risks that do not directly impact consumer safety are also not covered.

The research carried out by CSES involved a combination of desk research, an extensive interview programme and an online workshop with participants from all relevant stakeholder groups.

Terminology

Artificial intelligence (AI) is a broad term, defined by the Office for Artificial Intelligence and the Central Digital & Data Office as the "use of digital technology to create systems capable of performing tasks commonly thought to require intelligence". Although it is constantly evolving, this definition highlights that AI generally "involves machines using statistics to find patterns in large amounts of data" and has "the ability to perform repetitive tasks with data without the need for constant human guidance".[1] Machine learning (ML) – the field of study that gives computers the ability to learn "without being explicitly programmed"[2] – is a key subset of AI. AI involves **mimicking intelligent human behaviours**, such as learning, prediction and adaptability, often based on significant amounts of data.[3] This can manifest, amongst other things, as **pattern recognition,**

[1] UK Government's Central Digital & Data Office and Office for Artificial Intelligence. (2019). Guidance: A guide to using artificial intelligence in the public sector.
[2] Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers, IBM Journal of Research and Development 44:1.2 (1959): 210–229.
[3] OII & Google. (2020). *Artificial Intelligence*. The A-Z of AI. [online]

image recognition, optimisation, and recommendation generation based on data from a variety of media (videos, images, text, audio, etc.)[4]

However, **certain challenges persist regarding the use of the term artificial intelligence.** AI is often used as a buzzword in product marketing and is commonly conflated with related terms, such as 'smart' products, 'connected' products and consumer Internet of Things (IoT) products. As a result, the term AI is used to refer to a wide range of applications from quite simple algorithms to complex machine learning (ML) models.

Market for AI consumer products

In this context, it is **challenging to understand the true scale and dynamics of the market for AI powered consumer products.** With regard to the size of the market, quantitative data exists on the scale of the market for consumer IoT devices, robotics and the total market for AI. Although these data suggest a continuously growing market, they do not specifically provide information on the scale of the AI consumer product market. Qualitative data collected through interviews and supporting literature supports this general finding but indicates **notable differences between product groups.** While certain product groups, such as smart speakers, are found to be advanced in relation to the use of AI, other sectors, such as domestic appliances, report relatively limited use of AI in existing products.

Although the use of AI in consumer products is found to be increasing, the research found several **barriers to adoption**. Primarily, these include cost, privacy and awareness.

In light of the challenges related to terminology, it is possible to identify some **key characteristics of AI applications** that are relevant in a product safety context. In particular, AI (and primarily ML) systems often need significant amounts of good quality data for training, testing and validation purposes; and they can be opaque on two levels, as: (i) it is often not clear to a consumer when an AI system is in use; and (ii) the workings of the technical approaches themselves can be opaque. In addition, AI systems often have the ability to learn and develop over time, instead of relying on explicit instructions, and they can display autonomy in their actions and decision-making.

Opportunities and challenges for product safety

The **incorporation of AI systems into manufactured consumer products brings opportunities, as well as challenges and risks**. In terms of opportunities and benefits, there is a significant body of research highlighting the economic and social benefits of AI generally. When specifically considering product safety, the opportunities are also extensive, but can differ by product group. The direct opportunities for product safety include: more efficient and effective products; and predictive maintenance, which can directly improve product safety, as well as reduce maintenance costs and product downtime. In addition, indirect opportunities exist. These include: improved data collection and analysis in the different phases of industrial assembly to increase product quality; improved cyber security protection; AI powered product design; and increasing potential for personalised products.

[4] Ellen MacArthur Foundation. (n.d.). *Artificial Intelligence and the Circular Economy*. Ellen MacArthur Foundation. [online]

In relation to the challenges and risks of AI to product safety, **the characteristics of AI as a technology highlighted above (including mutability, opacity, data needs, and autonomy) can translate into errors or challenges for AI systems that have the potential to cause harm.** As illustrated in the figure below, these challenges can be categorised according to a range of themes, including robustness and predictability, transparency and explainability, security and resilience, fairness and discrimination, and privacy and data protection.

The **potential harms resulting from these challenges can be material or immaterial in nature**. Material harms, which are more likely to occur as a result of challenges in the first three themes (i.e. robustness and predictability, transparency and explainability, security and resilience), could include, for instance: an AI-driven robot malfunctioning as a result of automated decisions causing physical injury; or cyber security vulnerabilities in a product leading to threats to physical safety. Immaterial harms, which are more likely to occur as a result of fairness and discrimination or privacy and data protection challenges, could include, for instance, replacement of human contact for older people with autonomous products causing mental health issues; or discrimination in access to services for people with disabilities.

Beyond product safety risks specifically linked to AI, certain general trends can also bring product safety risks that can exacerbate or be exacerbated by AI consumer products. These include the tensions between built-in obsolescence and the circular economy, and the increasing reliance on e-commerce.

To date, however, **many of these risks are theoretical in nature and evidence of real-life examples of harm caused by AI consumer products is limited.** This most likely reflects a combination of factors, including: (i) the lack of maturity of many consumer product sectors in using AI; (ii) the existing consideration of the possible safety impacts of AI systems by the manufacturers and developers of these products; and (iii) the difficulty understanding the role and impact of AI systems when incidences do occur.

Beyond the potential impact of technical challenges on consumer harm, the **use of AI in consumer products can also challenge the regulatory framework for both product safety and liability**. For many existing AI consumer products, the current regulatory framework for product safety and liability and the mechanisms in place to monitor product safety are applicable and sufficient. However, the characteristics of more complex AI systems, in concert with general technological trends, pose challenges across all elements of the regulatory regime, including product safety and liability-related legislation, market surveillance regimes, standardisation, accreditation and conformity assessment. The key characteristics of AI systems, as highlighted above, include mutability, opacity, data needs and autonomy. The general market of trends of relevance include: the blurring of the lines between products and services; the increasing ability for consumer products to cause immaterial as well as material harm; the increasing complexity of the supply chains for

consumer products; and issues related to built-in obsolescence and maintenance throughout a product's lifecycle.

Considering the legal framework for product safety and liability, more complex AI systems, as well as general technological and market changes, challenge many of the definitions detailed by these laws. More specifically, it is not clear to what extent these developments fall within the existing definitions of product, producer and placing on the market, as well as the related concepts of safety, harm, damages, and defects. For instance, the definition of a product stipulated in the General Product Safety Regulations (GPSR) 2005 does not explicitly include or exclude software. Although coverage of software incorporated into a product before placement on the market is clearer, more significant challenges exist related to: (i) the coverage and impact of safety issues resulting from software downloaded on to, or third-party software incorporated into, a consumer product after it has been placed on the market; (ii) the coverage and impact of safety issues resulting from not maintaining software appropriately (i.e. through a lack of updates); and (iii) the coverage and impact of safety issues resulting from changes to a product after it has been placed on the market (e.g. through self-learning AI or software updates).

Furthermore, the characteristics of AI systems, the general trends highlighted, and the lack of clarity around the applicability of existing legal definitions and concepts, bring additional impacts. These include a lack of legal certainty for economic operators involved in the manufacture of AI driven consumer products, as well as a need to improve the skills and knowledge of regulatory bodies, such as MSAs and conformity assessment bodies, on AI systems.

1 Introduction

This document contains the Final Report for the 'Study on the Impact of Artificial Intelligence on Product Safety'. The assignment was conducted for the Office for Product Safety & Standards (OPSS) by the Centre for Strategy & Evaluation Services (CSES) between January and May 2021.

1.1 Study objectives and scope

The overarching objective of this exploratory study was to **examine the current and forecasted future impacts of artificial intelligence (AI) in consumer products, and what this means for product safety**. This breaks down into the following three specific objectives:

- **Objective 1:** Analyse the current and likely future applications of AI in the home, highlighting the advantages and disadvantages for consumers and product safety implications / risks.
- **Objective 2:** Assess whether the current product safety framework is sufficient for a new generation of products that incorporate AI.
- **Objective 3:** Examine what factors Regulators should consider when responding to these new challenges to ensure consumer safety and foster product innovation.

The following table provides an overview of our research framework, linking the three specific study objectives to research questions and signposting where in this report each issue is discussed.

Table 1-1: Summary of research framework

Study Objective	Research Questions	Final Report Section
Objective 1	What does the current and forecasted future market of AI-driven consumer products look like?	Market for AI consumer products analysed in section 3
	What are the key similarities and differences between AI and non-AI products and how can AI be defined in this context?	Distinguishing factors of AI analysed in section 2
	What new, inherent product safety impacts / risks / opportunities does the incorporation of AI present to consumers, businesses and regulators?	Challenges, risks and opportunities analysed in section 4
Objective 2	What is the existing product safety regulatory framework?	Descriptive overview provided in section 5

Study Objective	Research Questions	Final Report Section
	To what extent is the existing regulatory framework effective in ensuring product safety in AI-driven consumer products?	Regulatory opportunities and challenges are analysed in section 5
Objective 3	What characteristics of AI should be considered when regulating consumer products and how do these challenge existing product safety requirements?	Framework of policy considerations presented in section 6
	What possible policy options exist to respond to these product safety challenges, while facilitating and fostering product innovation, and what are their possible impacts?	

Considering the **study scope**, all manufactured consumer products subject to the General Product Safety Regulations (GPSR) 2005 and other relevant legislation for specific goods are covered (e.g. Electrical Equipment (Safety) Regulations 2016, Electromagnetic Compatibility Regulations 2016, Radio Equipment Regulations 2017 and Toys (Safety) Regulations 2011). Furthermore, it will be necessary to review literature across a range of geographies (UK, European, international), and consider the interplay between AI and other new technologies (e.g. IoT, 5G).

The scope of this study *does not include* vehicles, pharmaceuticals and food, nor does it cover risks associated with cyber security that do not directly impact consumer safety. More specifically, the cyber security vulnerabilities themselves and any impacts of the vulnerabilities that do not relate to product safety are out of scope.

1.2 Overview: Methodological approach

The data collection for this study included an extensive desk research exercise and interview programme. In total, interviews were conducted with 48 individual stakeholders including academics, research institutes, think tanks and tech hubs; government and public bodies; law firms; manufacturers, AI developers and industry associations; product safety practitioners and consumer associations; and standards bodies, notified bodies and testing labs.

The initial research covered the following topics: regulatory framework for product safety; relevant definitions and terminology related to the use of AI in consumer products; the current and future market for AI-driven consumer products; the opportunities and challenges related to the use of AI in consumer products; and the related regulatory opportunities and challenges.

Following the initial research, seven topics were selected in collaboration with OPSS for further in-depth research. These topics were: mutability; robustness and predictability; transparency and explainability; impact on vulnerable consumer groups; immaterial harm; existing approaches to tackling AI risks in consumer products; and liability. In addition, five short case studies would be developed, covering the use of AI in four specific product

groups (smart speakers, toys, robotics, white goods) and the challenges and opportunities brought by AI to market surveillance.

To validate the findings of the research, feedback was received from OPSS, as well as through a stakeholder workshop, a presentation to BEIS staff and peer-review by the Centre for Data Ethics and Innovation (CDEI) and the British Standards Institution (BSI). This report presents the findings across all research issues. Further information on the methodological approach can be found in Appendix C.

2 Understanding Artificial Intelligence

This section examines the key terminology related to Artificial Intelligence in the context of consumer products. It provides definitions of AI and machine learning, and explains what algorithms are and how they are utilised to produce outputs. It also provides an overview of the design and development process for machine learning models and highlights the distinguishing factors between AI and non-AI consumer products.

2.1 Terminology: Artificial intelligence and related terms

The term **Artificial Intelligence (AI)** was coined by John McCarthy in 1956 as "the science and engineering of making intelligent machines".[5] AI is a broad term referring to computer systems that can sense their environment, think, possibly learn and take action in response to what they are sensing or their objectives.[6] According to the Central Digital & Data Office and the Office for Artificial Intelligence, "AI can be defined as the use of digital technology to create systems capable of performing tasks commonly thought to require intelligence. AI is constantly evolving, but generally it:

- Involves machines using statistics to find patterns in large amounts of data;
- Is the ability to perform repetitive tasks with data without the need for constant human guidance".[7]

It refers to computer systems capable of tasks requiring some intelligence if performed by humans.[8] These tasks can be specific, also called 'weak' or 'narrow' AI (e.g. optimising electricity usage on a smart grid), or 'general' (e.g. an advanced chatbot).[9] AI systems are designed to operate with some degree of autonomy, however, they do not yet exhibit the same level of intelligence as a human. It should be noted that the scope and definition of AI is very contested, which makes it challenging to regulate.

Advances in AI have accelerated recently due to an evolution in machine learning and better computing power, data storage and communications networks. AI processes vast amounts of data, which might originate from various sources, such as images, video, sound or text, through software which draws conclusions, adjusts parameters and produces outputs. AI systems rely on algorithms to produce outputs. An algorithm is a set of instructions and operations, ranging from very simple to a very long and complex set of software and programming code, which processes the supplied data.

For example, the use of facial recognition to unlock a device, such as a mobile phone using Apple's Face ID, uses deep learning (see below) and works by placing 30,000 infrared dots on the user's face. It then uses algorithms to compare the user's face with the data it has stored to ascertain whether it is the genuine user trying to unlock the phone.[10] This form of facial recognition does not require a massive database of photos to determine

[5] McCarthy, J. (2007). What is Artificial Intelligence?
[6] PwC. (2018). The macroeconomic impact of artificial intelligence.
[7] https://www.gov.uk/government/publications/understanding-artificial-intelligence/a-guide-to-using-artificial-intelligence-in-the-public-sector
[8] Intellectual Property Office. (2019). Artificial Intelligence, A worldwide overview of AI patents and patenting by the UK AI sector.
[9] Intellectual Property Office. (2019). Artificial Intelligence, A worldwide overview of AI patents and patenting by the UK AI sector.
[10] Forbes. (2019). The 10 Best Examples Of How AI Is Already Used In Our Everyday Life.

an individual since it simply recognises one person as the user of the phone. Apple claims that the chance of a random person unlocking someone else's phone is around one in 1 million.[11] Moreover, Face ID is continually learning. Each time a phone is unlocked, it tracks small facial changes.[12] According to Apple, should there be a more significant change in the user's appearance, such as shaving off a full beard, Face ID confirms the user's identity with the passcode before updating the face data.[13]

Facial recognition technology is also used for other purposes, ranging from law enforcement to healthcare. To summarise, there are essentially four steps to facial recognition:

- Face detection;
- Face analysis – the software reads the geometry of a face;
- Converting the image to data – analogue information (a face) is turned into digital information (data). A faceprint is produced;
- Finding a match – the faceprint is compared against a database of other known faces.

Another example of the use of AI is smart speakers, which use speech recognition to ascertain requests by digitizing vocal sounds into a machine-readable format and analysing the words to determine what the consumer requires. A large amount of accurate, linguistic data needs to be incorporated into speech training to understand and respond to requests.[14] Speech recognition uses natural language processing (NLP) and deep learning neural networks (see below). "NLP is a field of artificial intelligence in which computers analyse, understand, and derive meaning from human language in a smart and useful way."[15] The software can then make determinations on what is being said, and then transcribes the words into text.

To provide an example, Alexa, Amazon's cloud-based voice service, works in the following way:[16]

- Amazon records words and sends the recording to Amazon's servers to be analysed;
- The requests are broken down into individual sounds and a database is consulted containing various pronunciations to determine which words correspond to the sounds;
- Important words are identified to understand the requests and carry out appropriate functions;
- Amazon's servers send the information back to the device and Alexa speaks.

[11] Kaspersky. What is Facial Recognition – Definition and Explanation.
[12] Las Vegas Review-Journal. (2017). Apple's Face ID technology can learn, but it takes time.
[13] Apple. About Face ID advanced technology.
[14] AI Business. (2020). How smart speakers work.
[15] Algorithmia. (2016). What is natural language processing? Introduction to NLP.
[16] Towards Data Science. (2018). How Amazon Alexa works? Your guide to Natural Language Processing (AI).

Machine learning (ML) is a subset of AI. In 1959, Arthur Samuel defined machine learning as the field of study that gives computers the ability to learn "without being explicitly programmed".[17] ML models are trained to find patterns in vast amounts of data to produce outputs based on new data. The two main approaches to machine learning are supervised and unsupervised learning. The below box presents an overview of these main approaches.

Box 2-1: Common learning approaches

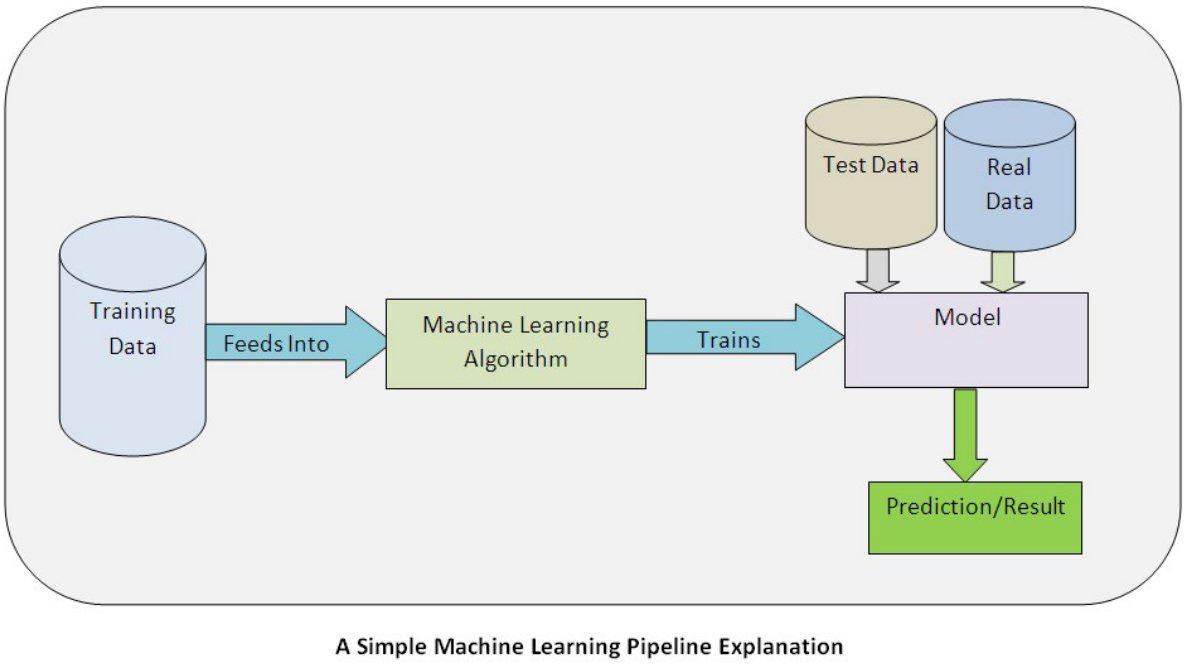
Supervised learning is a "learning strategy in which the correctness of acquired knowledge is tested through feedback from an external knowledge source."[18] In other words, data are labelled and the model can provide direct feedback on accuracy. For instance, in training a fruit classification model using supervised learning, a developer would label all the training data with the class name (i.e. apple or orange).[19] On this basis, the model can learn the characteristics of an apple and an orange (e.g. weight or texture) and therefore learn how to identify unlabelled data once deployed.[20] Although supervised learning models tend to be more accurate, they often require significant human resource up front to label the data appropriately.[21]

Unsupervised learning is a "learning strategy that consists in observing and analysing different entities and determining that some of their subsets can be grouped into certain classes, without any correctness test being performed on acquired knowledge through feedback from external knowledge sources"[22]. Continuing the above example, using unsupervised learning, the training data would not be labelled with the class name (i.e. apple or orange). Instead, the model would examine the characteristics of each fruit (e.g. weight, texture etc.), examine similarities and thus learn what constitutes an apple or orange without knowing what they are called.[23]

The below visual illustrates a simple development pipeline for a machine learning model. Further detail on the design and development process for ML models is presented in section 2.2.

[17] Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers, IBM Journal of Research and Development 44:1.2 (1959): 210–229.
[18] ISO/IEC 2382:2015(en) Information technology — Vocabulary
[19] Medium. (2019). 'Apple' or 'Orange' — Building Our First Machine Learning Model.
[20] Janarthanam, S. (2020). How do machines learn? Article on Medium.
[21] IBM. (2021). Supervised vs. Unsupervised learning: What's the Difference.
[22] ISO/IEC 2382:2015(en) Information technology — Vocabulary
[23] Janarthanam, S. (2020). How do machines learn? Article on Medium.

Figure 2-1: Simple machine learning pipeline explanation



Source: SQLServerCentral. (2019). Understanding Machine Learning.

With the advent of Big Data, machine learning has seen significant growth leading to the development of new algorithms. Access to vast amounts of specific data is crucial to the successful development of machine learning algorithms. Deep learning is a subset of machine learning that uses artificial neural networks to learn from vast amounts of data and perform tasks such as recognising speech or making predictions. Deep learning enables computer systems to perform and improve upon tasks by recognising data patterns. While ML uses simpler concepts, the artificial neural networks with which deep learning works are designed to imitate how humans think and learn.

Artificial neural networks are "an attempt to simulate the network of neurons that make up a human brain so that the computer will be able to learn things and make decisions in a humanlike manner. Artificial neural networks are created by programming regular computers to behave as though they are interconnected brain cells".[24]

Deep learning eliminates some of the data pre-processing that is typically involved with machine learning. The algorithms can process unstructured data and automate feature extraction, removing some of the dependency on humans. For example, if there was a set of photographs of different pets and categorisation by 'cat', 'dog' and 'hamster' was required, deep learning algorithms can determine which features (e.g. ears) are most important to distinguish each pet from one another (in ML, this is established manually by a human). The algorithm then adjusts and fits itself for accuracy, enabling it to make predictions about a new photograph with increased precision.[25] By observing patterns in the data, a deep learning model can learn about digital representations of text, image, sounds and other data and cluster inputs appropriately.[26] However, a deep learning model requires more data points to improve its accuracy, whereas a machine learning model

[24] Forbes. (2018). What Are Artificial Neural Networks - A Simple Explanation For Absolutely Anyone.
[25] https://www.ibm.com/cloud/learn/deep-learning
[26] The Consumer Goods Forum and IBM. (2019). Artificial Intelligence in Consumer Goods.

relies on less data given the underlying data structure.[27] Deep learning is primarily leveraged for more complex use cases, such as Face ID, virtual assistants (such as Alexa) and fraud detection.

Certain challenges persist regarding the use of the term AI and related terms, such as 'smart' products, connected products and the consumer IoT. AI is often conflated with these terms and used as a buzzword in product advertising. Principally, however, AI products mimic human behaviour and utilise machine learning and autonomy to arrive at decisions, which result from the processing of vast amounts of data and pattern recognition. Non-AI products do not behave in the same way and may simply be carrying out pre-programmed instructions. The case studies in this report will demonstrate the difference between AI and non-AI products but essentially, while AI makes decisions based on learning and the information it receives, automated products are pre-set and programmed to carry out a routine job. Additionally, automated products perform repetitive tasks based on commands and rules and do not evolve. One such example is automated responses when trying to book an appointment.

2.2 Machine learning: Design and development process

In the design and development of machine learning models, key actions and decisions need to be taken that can impact the success of an ML model and ultimately lead to product safety challenges. Although the specifics of the ML development process may differ by research team, there are common overarching actions that need to be considered:

Problem and goal definition. An important first step is to determine and define what problem the ML model intends to solve. For instance, two of the most common problem types are classification problems, where the model works to produce discrete output variables (e.g. labels or categories), and regression problems, where the model works to produce a continuous output variable (e.g. a quantity).[28] Once the problem has been defined, the goals of the system can be developed, and questions related to the types and availability of input data, the target features of the model and the expected outputs should be considered.[29]

Evaluation protocol development. Once the objectives of the model are clear, it is important to ensure achievement against the objectives can be measured and evaluated. To do this, the way success will be measured – i.e. the evaluation metrics – needs to be determined. The most appropriate evaluation metrics differ based on the problem being tackled. For instance, precision, accuracy, and recall are often used for classification problems, whereas using the mean squared error is common for regression problems. Following the selection of evaluation metrics, an evaluation protocol needs to be determined. This describes the process by which the evaluation of the model will be conducted and each method comes with its own set of trade-offs.

Data gathering, preparation and parsing. Within this step, the data necessary to train, test and validate the model needs to be collected, prepared, and separated. Once gathered, the raw data needs to be cleaned and transformed to ensure it is appropriate to facilitate efficient analysis and limit errors. Data preparation can include combining

[27] https://www.ibm.com/cloud/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks
[28] Brownlee, J. (2019). Difference between Classification and Regression in Machine Learning, Article in Machine Learning Mastery.
[29] Facebook. (2018). Blog: Introducing the Facebook Field Guide to Machine Learning, video series.

different data sets, dealing with missing data, making corrections, reformatting data, and standardising data formats.

Feature selection and scaling is a key aspect of this step. In machine learning, features are the individual independent variables used as inputs to the model.[30] For instance, in speech recognition, the features used can include noise ratios, length of sounds and relative power. The selection of appropriate features is considered to be a key indicator of ML model quality and, alongside model design, is one of the most important ways in which the performance of a model can be altered.[31] In addition, the selected features need to be scaled to ensure all features are on the same scale; this is commonly achieved through normalisation and standardisation.[32]

As mentioned previously, an ML model requires a significant amount of good quality and unbiased data to ensure its outputs are robust, accurate, fair, and representative. At this stage of the development process, it is therefore vital to assess biases in the data. Biases can exist in a data set for a range of reasons and can result in discriminatory or unfair outcomes. For instance, historical data may have codified human biases, or data sets may over- or under-represent certain demographic groups, data points or aspects of a phenomenon.[33] Due to high-profile examples of discriminatory outcomes by ML models, fairness issues have received significant academic and industry attention in recent years.[34] The impacts of fairness issues and the approaches developed to combat these issues are discussed in section 4.2.

Following preparation and pre-processing, the data is typically split into three sets, which have different purposes in the development process:[35]

- **Training set**: Used to train the parameters of the model.
- **Validation set:** Used to test the trained model and tune the hyper-parameters[36] of the model.
- **Test set**: Used to conduct an unbiased evaluation of the final model to determine its performance against the pre-defined success criteria.

Model creation, tuning and experimentation. This step requires the selection and tuning of a model, as well as comparisons between models. There are two main approaches to machine learning: supervised and unsupervised learning. Within each, a range of model choices exist. The choice of model needs to consider a range of issues, including the model's goals, the available data, interpretability, and ease to debug, as well as training and prediction considerations.

As mentioned above, a range of trade-offs exist in relation to model choice. More traditional, simpler models (such as linear models) tend to be more interpretable, easier to debug and work well with a wide range of data volumes, as compared with more complex

[30] Bishop, C. (2006). Pattern recognition and machine learning. Berlin: Springer.
[31] Facebook. (2018). Blog: Introducing the Facebook Field Guide to Machine Learning, video series.
[32] Roman, V. (2018). How To Develop a Machine Learning Model From Scratch, Article in towards data science.
[33] Veale, M. and Binns, R. (2017). Fairer Machine Learning in the Real World: Mitigating Discrimination without Collecting Sensitive Data, 4 Big Data & Society 205395171774353.
[34] Algo:aware. (2018). State-of-the-Art Report | Algorithmic decision-making. Report developed for DG CNECT.
[35] Facebook. (2018). Blog: Introducing the Facebook Field Guide to Machine Learning, video series.
[36] Hyperparameters are parameters that are used to control the learning process.

models (such as deep learning approaches). However, more complex models may be more appropriate when lots of training data and computing power is available.[37]

Machine learning models can also vary significantly in the extent of their settings. In this respect, developers need to consider the model hyperparameters (e.g. learning rate settings and regularisation choices[38]) and model architecture settings (e.g. feature interactions for linear models).[39]

Furthermore, infrastructure costs and capacity need to be considered. The features and models with the best performance may not be the best to implement as they may require more computing power than is available. This is particularly relevant for consumer products, which may have limited onboard computing power. As a result, consumer products are more likely to need to consider the trade-off between using the limited capacity to conduct ML work onboard and transferring data collected by a product to the cloud for analysis.[40]

As a result of the extensive range of trade-offs and options, a systematic and scientifically sceptical approach to experimentation, comparison of models and fine-tuning is needed. In the model training process, the key tension is between optimisation – achieving the best performance possible on training data – and generalisation – maximising performance on unseen data. The goal is to obtain the best generalisation ability without overfitting.[41,42] Model selection and tuning is another key point in the process where decisions can potentially impact product safety. Many research teams highlight a range of principles that need to be considered at this stage, including model efficiency, transparency, reproducibility, automation and comprehensiveness. The possible implications related to these principles will be discussed in more detail below.

Monitoring and maintenance. Once deployed, ML models require continuous monitoring and maintenance. In particular, issues such as model drift[43], opaqueness[44] and misdirected reinforcement learning behaviour[45] are common and can impact the performance of an ML model over time.[46] In addition to performance, ML models need to be monitored for errors, crashes and latency.[47] For consumer products, as highlighted by industry representatives interviewed for this study, over-the-air (OTA) updates and

---

[37] Facebook. (2018). Blog: Introducing the Facebook Field Guide to Machine Learning, video series.
[38] Regularisation is a technique used to ensure an ML model is appropriately fitted. Different types of regularisation techniques, such as ridge regression and Lasso, aim to do this in different ways. See more here: https://towardsdatascience.com/regularization-in-machine-learning-76441ddcf99a
[39] Facebook. (2018). Blog: Introducing the Facebook Field Guide to Machine Learning, video series.
[40] Facebook. (2018). Blog: Introducing the Facebook Field Guide to Machine Learning, video series.
[41] Roman, V. (2018). How To Develop a Machine Learning Model From Scratch, Article in towards data science.
[42] Overfitting refers to a machine learning model that "models the training data too well", thereby hindering application of the model to new data and negatively impacting a model's ability to generalise. See: Brownlee, J. (2016). Overfitting and Underfitting with Machine Learning Algorithms, Article in Machine Learning Mastery. [online]
[43] Model drift explains the performance decline that results from models being operated in dynamic environments with unfamiliar data.
[44] If a model is not working as intended, quality assurance to identify bugs is an approximative task and not easy in ML models. See: Schmelzer, R. (2020). 'Machine learning limitations marked by data demands'.
[45] European Commission. (2021). Study to Support an Impact Assessment of Regulatory Requirements for Artificial Intelligence in Europe, Final Report (D5)
[46] Appen. (2021). Blog: AI Model Maintenance: A Guide to Managing a Model Post-Production.
[47] Shin, T. (2020). Why You Need to Manage Your ML Models After Deployment, Article in towards data science.

upgrades for device software and firmware are vitally important to address these issues and to ensure ML models continue to perform effectively and efficiently.[48]

2.3 Distinguishing factors of AI vs. non-AI products

While AI is often conflated with smart, connected products or consumer IoT and used as a buzzword in product advertising, there are several key characteristics that define an AI product, and are truly innovative in a variety of consumer product markets. First, AI involves **mimicking intelligent human behaviours**, such as learning, prediction and adaptability based on the immense corpus of data it utilises, which can be gathered by the product itself or stored in vast quantities in databases.[49] This can manifest as **pattern recognition, image recognition, optimisation, and recommendation generation** based on data from a variety of media (videos, images, text, audio, etc.)[50]

Intelligence is markedly different from automation—when a device operates on basic, rules-based capabilities to replace repetitive manual and cognitive tasks carried out by humans. An automated device can simply accomplish tasks faster and perhaps at a larger scale than a human worker.[51] For example, a washing machine is programmed to clean a variety of clothes and other items via different settings for water temperature and time. Crucially, this device does not learn, and therefore does not require the gathering of vast amounts of data; it merely follows a pre-programmed protocol. A PWC report on AI's economic impact in the UK defines AI technologies as "computer systems that can sense their environment, think, learn and then take action as a result. This ability to respond to the environment stands artificial intelligence apart from automation of routine tasks."[52]

Such analytical capabilities are made possible by techniques such as **machine learning**, or the ability for a device to learn from the data it collects and make predictions based on its evolving understanding. Machine learning extends beyond simple image or pattern recognition, and can allow researchers to use such AI systems to understand immense quantities of data.[53] In consumer devices, machine learning already plays a critical role in, for example, facial recognition in security cameras. Connectivity both to other devices and, often, to the Internet means a device can constantly conduct pattern analysis to troubleshoot problems as they occur, predict national and local power grid failures before they occur, and learn which solutions are best to implement to prevent users from experiencing inconveniences or malfunctions.[54] Interviewees have expanded upon this feature, outlining how machine learning enables products to predict consumer needs, pre-empt their actions, and provide recommendations.

While **some non-AI products can collect data, they do not analyse it; rather, developers and other creators may examine the data offline in an ad-hoc analysis process.** It is therefore difficult at times to distinguish a non-AI from an AI product superficially, as the functionality does not necessarily change. Industry and academic interviewees have emphasised that at present, AI merely enhances a product's functions,

---

[48] Chauhan, A. S. (2020). A business case for Over-the-Air updates (OTA) in Smart Devices, Article in Becoming Human: Artificial Intelligence Magazine.
[49] OII & Google. (2020). *Artificial Intelligence*. The A-Z of AI. [online]
[50] Ellen MacArthur Foundation. (n.d.). *Artificial Intelligence and the Circular Economy*. Ellen MacArthur Foundation. [online]
[51] PWC. (2017). *The economic impact of artificial intelligence on the UK economy*. PWC.
[52] PWC. (2017). *The economic impact of artificial intelligence on the UK economy*. PWC.
[53] OII & Google. (2020). Machine Learning. The A-Z of AI. [online]
[54] AT&T Foundry, Ericsson, & Rocketspace. (2018). *The Future of Artificial Intelligence in Consumer Experience: According to the AT&T Foundry*.

perhaps with some performance optimisation. This will be especially important to bear in mind as uptake of IoT and smart products in the home increases.

3 Market for AI consumer products

This section examines the market for AI-driven consumer products and the AI applications currently in use in consumer products. More specifically, the section covers: (i) the evolution and scale of use of AI across consumer product groups; (ii) the types of AI applications in use, considering both user functionality and algorithmic approaches; and (iii) the future development of the market.

3.1 AI applications in consumer products

Below we provide an overview of the current market for AI consumer products in the UK and wider afield. As mentioned previously, the term AI is often conflated with other products, such as smart products. Whilst there are differences between AI and non-AI products, as outlined in Section 2.3, the products covered in this section often simply fall under the category of 'smart'. As such, it is highly likely that the analysis provided will cover both AI and non-AI consumer products, as the statistics available do not often differentiate between these types of products. In summary, the **use of smart products** is growing both in the UK (and internationally).

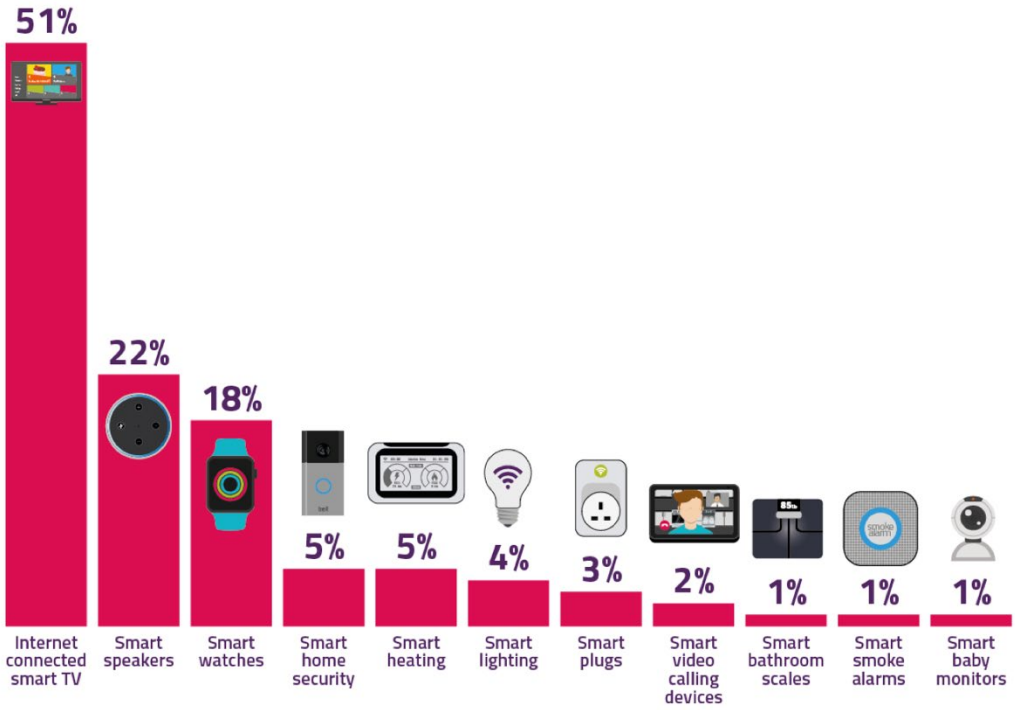
Indeed, when asked about the market, interviewees from all stakeholder groups were generally of the opinion that **the term AI is used far more than its actual application**, whereby there are products claiming to use AI but do not. As such, while consumers are certainly increasing their usage of products with AI, as will be seen, it is possible that the actual AI market is smaller than perceived.

According to a recent report published by Ofcom, a third of adults have at least one smart technology device, excluding smartphones and smart TVs,[55] among the devices covered in Figure 3-1 below. Over 50% have an internet-connected smart TV, while 22% own a smart speaker. Furthermore, according to the Ofcom report, **11% of all UK households own some kind of 'smart home' technology** (a sub-category that includes devices such as smart home security, smart lighting and smart heating). Research by YouGov from 2018 has shown that 23% of Britons own at least one smart home device (including smart speakers, but excluding smart meters), with 8% owning two or more.[56]

This trend is driven primarily by **increases in the number and type of consumer Internet of Things (IoT) devices**, some of which incorporate AI systems. Many smart speakers, for example, integrate AI-driven speech recognition and voice assistant systems to understand and respond to user requests, though quantitative data on the use of these specific features is lacking. Convenience, fun and enjoyment of trying new technology and more, better features than non-internet connected options have been cited as reasons for having smart technology devices in the household.

[55] Ofcom. (2020). Online Nation: 2020 Report.
[56] YouGov. (2018). Almost a quarter of Britons now own one or more smart home devices.

Figure 3-1: Prevalence of smart technology devices in UK households (2020)



Source: Ofcom. (2020). Online Nation: 2020 Report.

The burgeoning nature of the market can be illustrated by the recent investments in IoT made by tech giants, such as Google, Apple, Amazon and Alibaba, which has transformed the market. Google is by far the biggest investor of these, having spent USD 3.9 billion (GBP 2.6 billion) acquiring AI startups since 2006. Amazon has spent USD 871 million (GBP 626 million acquiring startups, while Apple has spent USD 786 million (GBP 565 million). Intel and Microsoft complete the top 5, having invested USD 776 million (GBP 558 million) and USD 690 million (GBP 496 million) respectively. As can be seen, the market is still driven by the tech giants and will likely be in the near future.[57] That said, there are numerous medium-sized companies producing AI products, and AI startups raised USD 33 billion (GBP 24 billion) in 2020.[58]Consumers are becoming more and more familiar with smart home devices. A survey conducted by techUK, for instance, found that **nearly 80% of consumers know at least something about smart homes.**[59] The growth in the smart product market is also due to collaboration between smart home manufacturers and developers of smart speakers, enabling the integration of smart home devices with smart speakers.[60]

While the use of smart home products is certainly on the rise, cost, privacy and awareness have been cited as barriers to adoption. Indeed, privacy concerns are particularly apparent in the smart entertainment category, while there has been much discussion over whether smart speakers are picking up conversations and the extent to which providers conduct manual quality checks to improve voice recognition.[61] **Equally, manufacturers may not be willing to take the risk to incorporate AI into consumer products.** For example, as

[57] TechRepublic. (2018). The 10 tech companies that have invested the most money in AI.
[58] https://builtin.com/artificial-intelligence/ai-companies-roundup
[59] techUK. (2020). The State of the Connected Home.
[60] Kumar, R., & Rasal, A. (2018). Smart Speaker Market by Intelligent Virtual Assistant, End User, Distribution Channel, and Price – Global Opportunity Analysis and Industry Forecast, 2018-2025.
[61] techUK. (2020). The State of the Connected Home.

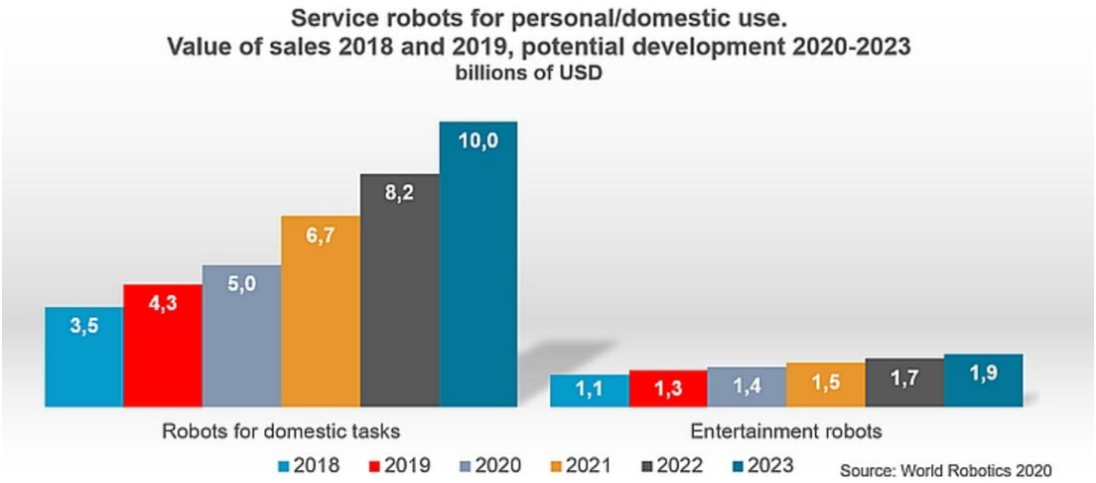
one industry interviewee pointed out, the toy sector generally withdrew from developing such products several years ago due to issues with pioneering products, meaning today there is very little AI used in toy products. According to the interviewee, the market has not grown, and it is unlikely it will anytime soon. However, this lack of growth appears to be an anomaly rather than the general trend.

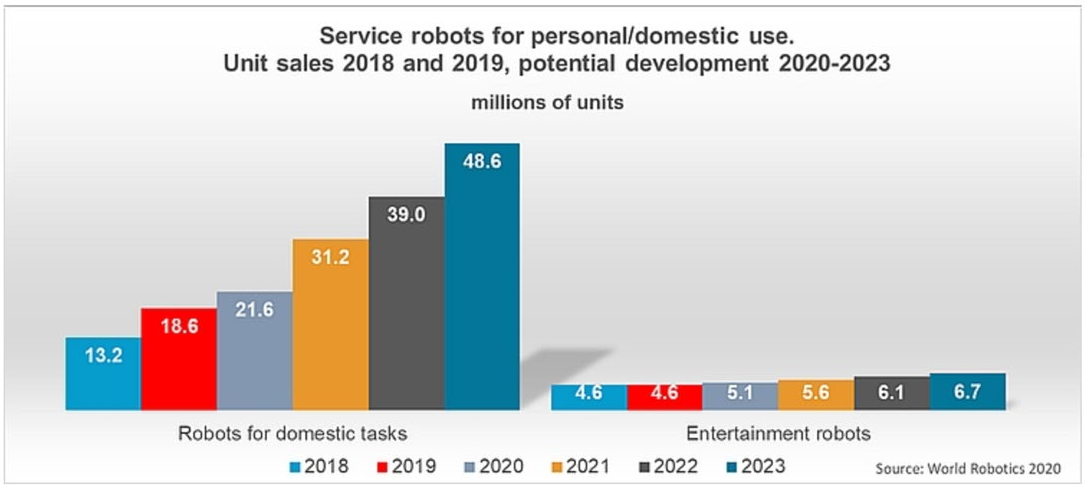
Research conducted for DCMS in 2020 analysed key evidence related to the future evolution and size of the consumer IoT market and the findings further illustrate this trend of growth.[62] For instance, in 2017, Ofcom predicted that by 2024, **the UK would have around 156 million IoT connections, increasing from around 13 million in 2016.**[63] Key consumer IoT market segments, including Consumer Electronics and Fast-Moving Consumer Goods (FMCG) (39.9 million) and Utilities (36.5 million), will account for nearly half (76 million) of these future connections.[64] Globally, this figure has been estimated at 500 billion connected devices by 2020.[65] It has also been suggested that in 2017, the global smart home market was estimated to be worth USD 43.4 billion (GBP 31.4 billion)[66] and has more than doubled since then, with recent projections showing that this figure was expected to reach USD 91 billion (GBP 65.8 billion) in 2020.[67] The global market is expected to have an annual growth rate of 15% by 2024.[68]

The global market for personal and domestic service robots is also following this trend. **According to the International Federation of Robotics (IFR), over 23.2 million units were sold in 2019, an increase of 34%, resulting in global sales of USD 5.7 billion (GBP 4.1 billion).**[69] By 2023, this figure is predicted to more than double to USD 12.1 billion (GBP 8.7 billion),[70] with an estimated 55.3 million units being sold.[71] These robots include vacuuming and floor cleaning robots, lawn-mowing robots and entertainment robots. The figures below demonstrate the sales and estimated sales until 2023. The increase can partly be attributed to falling prices, with the unit price of robot vacuums and toy robots having declined in recent years. For instance, robot vacuums can now be purchased for less than USD 100 (GBP 72.3).[72] [73] The use of assistance robots for elderly or handicapped persons is also on the increase, with estimated sales valuing USD 91 million (GBP 65.8 million).[74]

[62] CSES, (2020), Framing the Nature and Scale of Cyber Security Vulnerabilities within the Current consumer Internet of Things (IoT) Landscape, study for DCMS.
[63] Cambridge Consultants, (2017), Review of the latest developments in the Internet of Things, study for Ofcom.
[64] Ofcom, (2017), Connected Nations 2017: Data analysis, pp. 47-49.
[65] European Commission, (2020), Combined Evaluation Roadmap/Inception Impact Assessment: GPSD and AI.
[66] As of 08/03/2020, using a conversion rate of 1.38.
[67] As of 08/03/2020, using a conversion rate of 1.38.
[68] techUK. (2020). The State of the Connected Home.
[69] As of 08/03/2020, using a conversion rate of 1.38.
[70] As of 08/03/2020, using a conversion rate of 1.38.
[71] International Federation of Robotics. (2020). World Robotics - Service Robots 2020.
[72] IFR. (2020). Service Robots Record: Sales Worldwide Up 32%.
[73] As of 08/03/2020, using a conversion rate of 1.38.
[74] As of 08/03/2020, using a conversion rate of 1.38.

Figure 3-2: Sales and estimated future sales of service robots for personal/domestic use



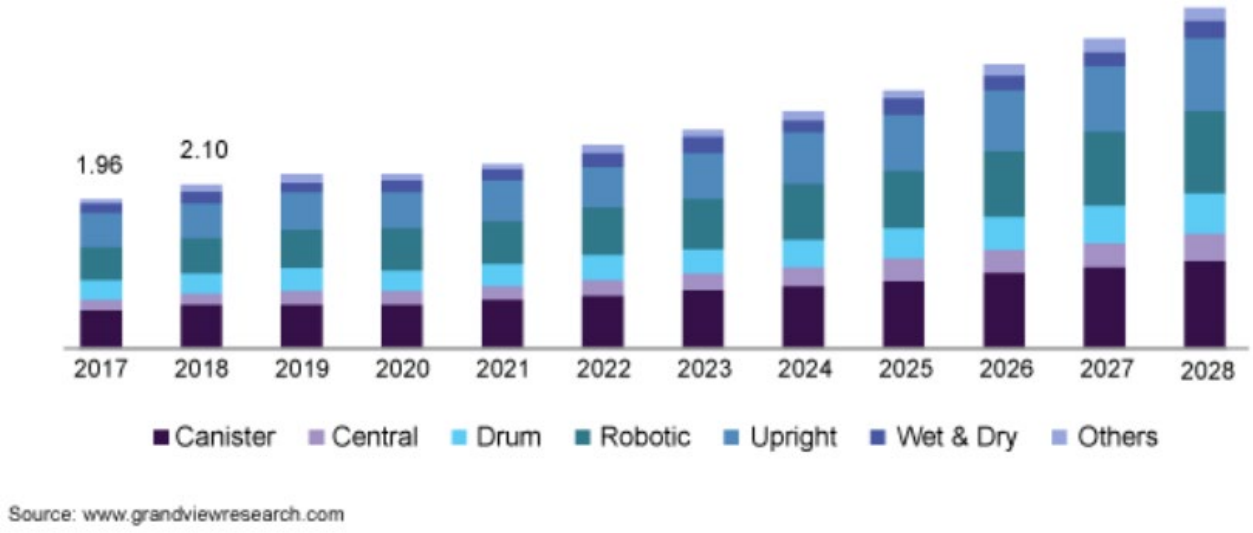


Source: IFR. (2020). Service Robots Record: Sales Worldwide Up 32%.[75]

To put these figures into context, research has shown that the global market for vacuum cleaners as a whole, including robots, was estimated to be worth USD 10.01 billion (GBP 7.25 billion)in 2020 and is expected to increase at a Compound Annual Growth Rate (CAGR) of 9.6% from 2021 to 2028.[76] Considering robotic vacuum cleaners specifically, other market research estimates that the global market was USD 2 billion (GBP 1.45 billion) in 2020 and will be worth USD 3.4 billion (GBP 2.4 billion) by 2027.[77] Figure 3-3 below shows the market size proportion of robotic vacuum cleaners in the US compared to other vacuum cleaners.

[75] IFR. (2020). Service Robots Record: Sales Worldwide Up 32%.
[76] Grand View Research. (2019). Vacuum Cleaner Market Size, Share & Trends Analysis Report.
[77] GlobeNewswire. (2021). Global Robotic Vacuum Cleaners Industry.

Figure 3-3: US vacuum cleaner market size, by product, 2017 – 2028 (USD billion)



Source: Grand View Research. 2021. Vacuum Cleaner Market Size, Share & Trends Analysis Report.

Research undertaken by the same company has shown that the global lawn mowers market size was USD 28.5 billion (GBP 20.5 billion) in 2019.[78] The electric segment held the largest market share of 29.7% in 2019, while the robotics segment is expected to witness the highest CAGR of more than 11% over the forecast period, compared to manual, electric, petrol and others. These figures demonstrate that, while the use of robots for cleaning and lawn mowing is certainly increasing, consumers still prefer conventional machines. This is unsurprising, however, given that these products were only brought to the market in recent years. The use of AI in robotic vacuum cleaners and lawn mowers is explored further in a case study on the use of AI in robotics (Appendix A).

It has also been suggested that **hardware improvements will accelerate the use of domestic robots over the next few years.** Better chips, low-cost 3D sensors, cloud-based machine learning and advances in speech recognition will improve the services provided by robots as well as their interaction with humans.[79] One interviewee from the legal industry mentioned that as AI solutions become more sophisticated, the activities they are involved in will increase. In the coming years, AI is expected to transform several sectors, including:

- Transport: self-driving vehicles are expected to be widely adopted. These include cars, delivery trucks, autonomous delivery drones and personal robots.
- Service robots: these are expected to deliver packages, clean offices and enhance security.
- Healthcare: patient monitoring, clinical decision support and surgery assistance are some of the potential applications of AI in healthcare.
- Education: AI can provide personalisation at scale. NLP, ML and crowdsourcing can be integrated with face-to-face learning to enhance the learner's experience.

[78] Grand View Research. (2020). Lawn Mowers Market Size, Share & Trends Analysis Report.
[79] Stanford University. (2016). Artificial Intelligence and Life in 2030.

- Entertainment: AI-enabled entertainment is expected to become more interactive, personalised and engaging by 2030.[80]

The increase in the use of consumer IoT products reflects the growing global AI market, which has the potential to bring about **major economic and social benefits.** One estimate suggests that the worldwide market for AI solutions could be valued at more than GBP 30 billion by 2024, increasing productivity by up to 30% in some industries and generating savings of up to 25%. Another estimate suggests that AI could contribute up to USD 15.7 trillion (GBP 11.3 trillion)[81] to the global economy in 2030, of which USD 6.6 trillion (GBP 4.8 trillion)[82] is likely to come from increased productivity and USD 9.1 trillion (GBP 6.6 trillion)[83] from consumption-side effects.[84]

Another report by McKinsey Global Institute (MGI) predicts that **AI could increase global output by approximately 16% by 2030, noting that this figure could be marginally larger, at 22%, for the United Kingdom since it is potentially more AI-ready compared to the global average.**[85] The report suggests that this growth in GDP would be a welcome increase for the UK, considering its recent weak performance in productivity growth compared to other advanced economies. Between 2010 and 2015, productivity increased by only 0.2% per year, whereas the average between 1970 and 2007 stood at 2.4%. However, the report recognises that any growth emanating from AI adoption will take time due to lags and transition costs, including accessing talent. In another report, PWC predicted that UK GDP will be up to 10.3% (GBP 232 billion) higher by 2030 due to the incorporation of AI.[86] The report attributes the UK's AI-readiness to seven enablers: research; start-up investment; automation potential; digital absorption; innovation foundation; human capital; and ICT connectedness. The UK is currently in the top quartile of the MGI index which includes China, the US and the EU Member States.

The drive in the use of AI will also be facilitated by widespread adoption of 5G, which has already been introduced in the UK. 5G provides the infrastructure and vast amounts of data AI needs to enhance productivity on an industrial scale while also benefiting consumers through more intelligent products which can save consumers time and money. Various factors will slow or accelerate the deployment of AI; these include understanding and testing AI, bringing it to the market, taking a customer-centric, data-driven approach, and partnerships and co-innovation.[87] AI has the potential to benefit different sectors, including healthcare, transport, environment, farming and smart cities, which will increase the development of AI technologies.

The way in which AI is applied is dependent on the individual product. Smart heating works by using algorithms to save the consumer energy based on the consumer's behaviour and schedule. The Nest Learning Thermostat, for example, uses an ML model that learns the consumer's ideal temperature and programs itself in around a week. The thermostat can also automatically turn the heating off to save energy if the consumer is not at home. Voice assistants use machine learning to better understand natural language questions and requests. Smart home security devices recognise faces and detect suspicious sounds to alert home owners of intruders. Entertainment robots can recognise

[80] Fingent. (2020). How Will Artificial Intelligence Transform The World By 2030.
[81] As of 08/03/2020, using a conversion rate of 1.38.
[82] As of 08/03/2020, using a conversion rate of 1.38.
[83] As of 08/03/2020, using a conversion rate of 1.38.
[84] Hall, W., & Pesenti, J. (2017). Growing the artificial intelligence industry in the UK.
[85] McKinsey Global Institute. (2019). Artificial intelligence in the United Kingdom: Prospects and challenges.
[86] PwC. (2017). The economic impact of artificial intelligence on the UK economy.
[87] GSMA. (2019). AI & Automation: An Overview.

and remember people using an in-built camera. Although many consumer IoT products incorporate AI to improve the products and related services, this is not true for all consumer IoT products. Consumer products that are not connected to the internet may also integrate AI systems. For example, a robot vacuum does not need Wi-Fi connectivity to sense its surroundings, follow its in-built hard-coded rules and navigate a room. In recent tech developments, there is the potential to gradually wean AI technologies off memory, computational, or internet and cloud sources. In these situations, the device's system adapts to the lack of resources via its deep learning capabilities in order to continue functioning. While this area requires further research, it has the potential of reducing the energy AI systems consume, as well as the size of devices.[88]

3.2 Sectors developing AI products

There are already clear ways in which AI can enhance consumer products by increasing their quality, increasing consumer choice through more personalised, varied goods, and saving consumers time by being able to multitask better and delegate to AI technologies.[89] In this description, users are still at the centre of their smart product environment, and in control of how their devices are used. Optimisation of device functionality, based on user data and feedback, allows for the increase in quality to meet their needs. In the coming years, this could improve consumer trust in AI, which has been lacking, and encourage innovation across various sectors to improve different domestic and lifestyle services.[90] According to a consumer survey conducted by European consumer organisation BEUC, the vast majority of respondents consider AI to be somewhat or very useful, particularly when used to predict traffic accidents (91%), their health (87%) or financial problems (81%).[91] However, there is a significant lack of trust, with a large majority having medium or low trust in privacy protection when using AI devices, with wearables and voice/virtual assistants highlighted in particular.

Alongside the increase in popularity—both among manufacturers integrating AI capabilities in consumer products, and consumers who purchase such products—there are a number of different sectors developing and selling consumer AI products. As mentioned previously, large technology companies are leading the way and have spent billions developing AI solutions and capabilities. This spending has eliminated potential rivals and concentrated capabilities among a few companies.[92] The remainder of this section highlights some of the key sectors in which AI use in consumer products is prominent. The following table provides an overview of the sectors, products and types of AI in use.

Table 3-1: Overview of key sectors developing AI consumer products

Sector	Product(s)	Type of AI
Entertainment	Mobile phones, tablets, computers, TVs, wearables	Voice recognition, natural language processing, image recognition, machine learning

[88] Nelson, P. (2017). Artificial intelligence may not need networks at all. Networkworld.
[89] PWC. (2017). *The economic impact of artificial intelligence on the UK economy*. PWC.
[90] techUK. (2020). The State of the Connected Home
[91] BEUC. (2020). Artificial Intelligence: what consumers say.
[92] Bloomberg. (2020). Big Tech Swallows Most of the Hot AI Startups.

Communication	Mobile phones, tablets, computers	Voice recognition, natural language processing, image recognition, machine learning
Domestic appliances	White goods and other household appliances	Voice recognition, natural language processing, image recognition, machine learning
Speaker & Soundsystem	Speakers, smart assistants	Voice recognition, natural language processing, machine learning
Energy & Gas	Smart meters and thermostats	Machine learning
Security	Doorbells, security cameras (sold in kits), alarms	Facial and image recognition
Independent (miscellaneous)	Wearables, biometric AI devices, vacuum cleaners, toys	Machine learning

First, there are existing **mobile phone and technology** manufacturers, such as Apple, Google and Samsung. These are companies that established themselves through software development, and / or mobile phone and desktop computer and laptop models. Alongside the boom in smart consumer product uptake, these manufacturers may have broadened their product offering to include TVs, wearables, and tablets, all equipped with voice assistants and machine learning capabilities. Indeed, these high-tech and telecoms firms are looking at incorporating more advanced forms of AI in their products, with 30% of respondents from this sector reporting their companies have embedded deep-learning capabilities in their products and services.[93] These companies benefit from established brand recognition among consumers, which can aid in adoption of new products if consumers deem these brands reliable.

An offshoot of this sector covers established **speaker and sound system** manufacturers. Such companies may have historically offered a variety of products, such as musical instruments, radio or home sound systems, and now integrate AI capabilities into their products as an enhancement of the features they already provided.[94] Learning from user requests can enable playlist creation, song recommendations, or enable hands-free audio playback in the home. As with other smart technologies, these speakers are often equipped with Internet connectivity or a companion smartphone app, which allows users to ask questions via voice command and control their devices when they are connected to the Internet. However, the device control feature is not an AI capability.

Third, smart meters are not only provided by independent manufacturers, but also established **energy and gas** companies. Again, this is an example of familiar, experienced companies adopting the latest technology to meet consumer demands and evolving trends in their sector, and applies to both large and small energy suppliers. AI is used to analyse data from smart meters with machine learning to understand energy use and predict patterns of consumption in the future, ensuring security of supply.[95] As of 30 September 2020, 22.2 million smart and advanced meters had been installed in homes and small

[93] McKinsey. (2020). *The state of AI in 2020*. McKinsey & Company. [online]
[94] Which? (n.d.). Wireless, smart and Bluetooth speakers. Which? [online]
[95] Apolitical. (2020). The risks and benefits of AI smart meters.

businesses in the UK.[96] The same can be said for **security companies**, which provide products such as cameras, doorbells, alarms, and other devices for both home security and commercial properties.[97] These are typically sold to consumers as bundles or kits for installation throughout the home.

Finally, there are **independent AI product manufacturers**, who often disrupt and lead the market in terms of innovation by creating products that use AI in unique ways, begetting a number of imitators and competitors from both established and independent companies. Fitness tracker wearables, other wearable biometric monitoring devices (BMDs), are one such example. Wearables are equipped with sensors that gather a preponderance of data on their users' physiological or biological activity (such as heart rate, steps taken, blood pressure, sleep patterns, etc.), presumably with greater detail and granularity than traditional clinical observation.[98] This data is then aggregated and presented to users alongside inferences or predictions either via the device or companion apps, which can both advise clinicians in addressing patients' needs and allow all types of users to keep an eye on their physical wellbeing. Products can range from smartwatches to AI-enabled skin screening services[99] to smart shirts.[100] Major players in the smartwatch market include Apple, Samsung, Garmin and Fitbit. According to research, companies are prioritising R&D investments to increase their market shares. Additionally, the leading companies are collaborating with suppliers and resellers to strengthen their market positions through selecting the right channels, regions and target audiences.[101]

Other independent, product-specific manufacturers include those for **smart vacuum cleaners**, smart washing machines and smart fridges, which will be explored in the case studies. Smart vacuum cleaners clean floors without human contact guiding them, utilising an intelligence system linked to multiple sensors, such as wall, cliff, or object sensors to learn how to navigate a user's home more precisely.[102] Academic interviewees have mentioned how, in the future, such products could use machine learning to predict how long it takes for their bag to fill up, and notify its owner about when it is time to replace the bag, or learn when to notify its owner about maintenance matters based on data from other homes' smart cleaners. While other innovators have been inspired by this product category, such as developing a smart carpet that directs the cleaner around the room, these products do not necessarily use AI.[103] There is potential for these smaller innovations to adopt AI in the future, but at present that is not the case.

Some industries have yet to digitalise. A global 2019 report notes that **50% of manufacturers still rely on analogue processes, such as Excel spreadsheets for inventory tracking, and around 58% supply management sources have yet to adopt a centralised data storage system**, which could provide valuable insight on market

[96] BEIS. (2020). Smart Meter Statistics in Great Britain: Quarterly Report to end September 2020. BEIS.
[97] Which? (n.d.). *Smart home security systems*. Which? [online]
[98] Arnerić, S. P. et al. (2017). Biometric monitoring devices for assessing end points in clinical trials: developing an ecosystem. *Nature Reviews Drug Discovery*. 16(736).
[99] Esteva, A. et al. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature 542, 115–118.
[100] Bobin, M., Amroun, H., Anastassova, M., Boukallel, M. & Ammi, M. (2017). In IEEE International Conference on Systems, Man, and Cybernetics.
[101] Allied Market Research. (2020). Smartwatch Market, 2020-2027.
[102] Layton, J. (n.d.). *How Robotic Vacuums Work*. Howstuffworks. [online]
[103] Layton, J. (n.d.). How Robotic Vacuums Work. Howstuffworks. [online]

behaviour.[104] The main barrier to these and other adoptions is cost, as 68% of respondents preferred to save money rather than take a risk with innovation.

Academic interviewees have predicted that **within the next 10 years, there may be entirely new product categories, which operate autonomously and flexibly** (changing and altering their processes throughout their operation, rather than being turned to a certain setting). In addition, there may be higher volumes of interactions between devices within the home. One example provided was training an AI smartphone assistant to better recognise the nuances in a user's speech, which simultaneously improved the speech recognition accuracy of another smart product owned by the user.

However, these interviewees and those within government have indicated that **AI development is, in some ways, approaching a plateau**; consumers are more informed about the existential and moral risks of using historical data (such as embedding biases from those past datasets into a system, which leads to biased analysis methods and conclusions/recommendations), and product developers have yet to figure out how to overcome that. However, the evidence supporting a definitive plateau is limited; rather, these factors can be viewed as slowing the development of AI, and in time it will become clear whether innovation can work through these limiting factors. A survey conducted by BEUC found that consumers are concerned about privacy protection, AI manipulating their decisions, the reliability and safety of AI and the allocation of responsibility and liability if there is a problem, in addition to bias.[105] Without a better understanding of how to create products that benefit all consumers, and not just members of a few specific demographic categories, progress in further developing AI capabilities may slow down. These interviewees also mentioned that voice command technology will become further developed and integrated into more products.

[104] Bayern, M. (2019). Manufacturers' digital transformation initiatives lag behind other industries. TechRepublic [online]
[105] BEUC. (2020). Artificial Intelligence: what consumers say.

4 AI consumer products: Product safety opportunities and risks

The **incorporation of AI systems into manufactured consumer products brings new opportunities, as well as challenges and risks**, some of which are relevant from a product safety perspective and may necessitate Government intervention. This section first assesses the product safety opportunities, before considering the risks and challenges.

4.1 Opportunities and benefits

The primary value of AI systems is **their ability to perform complex analytical tasks in real-time** that would not be possible for humans (e.g. identifying patterns or processing vast amounts of data). This ability can deliver positive product safety impacts throughout the value chain. To date, despite the advances in AI and its widening availability, these advantages have been more widely recognised at the manufacturer and manufacturing process level and less so at the consumer level. However, the **application of AI to consumer products can lead to enhanced safety outcomes for consumers**. This can happen on an indirect basis, whereby consumers benefit from enhanced product safety performances through AI-led improvements in their manufacturing processes, or on a direct basis where consumers could benefit directly from the embedding of AI in products, which could identify unsafe product usage or optimise product performance.

Among the **direct advantages**, one of the main opportunities of AI in consumer products lies in leveraging sensors to provide real-life insights on how a product is being used by consumers and performing. This can give critical information to manufacturers on when a product embedded with AI might need repairs. The ability to predict repair needs is called predictive maintenance, enabling organisations to forecast when equipment will fail so that its repair can be scheduled on time. By establishing when an intervention is needed, predictive maintenance can play a key role in preventing accidents from occurring due to malfunctioning or product failure. Although outside of the scope of this study, examples of predictive maintenance have already been implemented in the automotive sector[106], who use car data to predict failures and servicing needs and enhance monitoring, allowing them to improve their safety records. Moreover, through preventing product malfunction, predictive maintenance allows for downtime reduction and could increase a product's lifetime while also reducing maintenance costs.

Below are a few examples of how the application of **AI can provide direct and tangible safety benefits to consumers**:

- Headphones may use AI to ensure a better experience for consumers and ensure hearing health longer-term.[107] Moreover, a patent has been published which introduces AI technology that can tell headset wearers when they are in danger by alerting them of imminent risks, such as a speeding car which might not be heard by the users of standard headsets.[108]

[106] IBM. (2016). How content analytics helps manufacturers improve product safety and save lives.
[107] Daudu, A. (n.d.). These Headphones use Artificial Intelligence to protect your hearing.
[108] Electronic Product Design and Test. (2019). Headphones gain AI-powered 'sense of hearing' via new Audio Analytic patent.

- Refrigerators are being equipped with AI to maximise their lifespan. South Korean manufacturer LG is introducing AI systems that can alert on refrigerator misuse, malfunction, or if they need maintenance while also informing if the fridge is overloaded, has limited airflow, or presents temperature fluctuations.[109] This is operated through its ThinQ app which uses deep learning.

- Although outside the scope of this study, some automotive drivers are already benefitting from the enhanced protection afforded by AI, which reminds drivers to follow specific traffic rules, maintain a safe distance between other vehicles and drive in the correct lane.[110]

Regarding indirect benefits to consumers' safety, AI can **improve product safety by enhancing the data collection processes during the industrial assembly** to prevent mass product recalls. This can be achieved through a 'digital twin' system that uses computerised versions of different physical assets connected via multiple sensors, which report on positions, key metrics, and physical issues on a more granular level. This type of AI can detect problems that are not observed through manual inspections and enable issue-detection before products are sold. This allows the discovery of 'rare events' which don't follow the routine patterns occurring on the assembly line by modelling each physical asset and workflow individually, identifying every possible situation that could lead to an issue. Humans cannot do this as some components have millions of assets that cannot all be individually modelled through manual data inputs.[111] The cognitive technology firm, DataRPM uses digital twin algorithms to teach machines to detect failures and quality issues on the assembly line. Through the use of 'digital twins', a producer might be able to source alternative suppliers in the case of supply chain dislocations, as happened during the Covid-19 pandemic.[112] Virtual agents can deliver instructions on tablets or other devices to reduce assembly errors and teach new operators in the workplace.[113]

Other tools such as visual recognition can assist the processes involved in conducting quality inspections along the supply chain and ensuring the quality of all the parts and components being assembled.[114] AI can continue to contribute to product safety even after a product leaves the production line, where it can be used to track products' quality performance post-distribution. Moreover, researchers in the US have demonstrated that businesses could use AI to identify products that need to be recalled partly by analysing Amazon's product reviews.[115] In addition, by helping ensure the functioning of critical infrastructures, such as energy utilities, AI can support safe and adequate energy supplies for producers.[116]

AI can also potentially contribute to product safety at the design stage.[117] Indeed, **AI may aid the tasks fulfilled by engineers and other professionals** by allowing them to input information on restrictions, production methods, material and other variables into an algorithm that can cut their time and effort. The algorithm may then develop only solutions that are safe, allowing designers and engineers to focus on the design aspects. The use of

[109] Gebhart, A. (2019). LG's new AI will tell you if you're using your fridge wrong.
[110] Majewsky, S. (2020). Can Artificial Intelligence Make Us Safer?
[111] CloudTweaks. (2016). Digital Twin and the End of the Dreaded Product Recall.
[112] CSES. (2020). Opportunities of Artificial Intelligence.
[113] McKinsey. (2017). Artificial Intelligence the Next Digital Frontier?
[114] IBM. (2019). Artificial Intelligence in Consumer Goods.
[115] Consumer Affairs. (2019). The future of product recalls: AI and Amazon.
[116] Bilodeau, S. (2019). Artificial intelligence in a "no choice but to get it smart" energy industry!
[117] Becominghuman. (2019). How AI Can Improve Product Safety.

AI in customer service can also contribute to product safety through virtual assistants, which can answer queries and provide recommendations on safe usage. In the future, it might be possible that even more complex tasks such as those performed by doctors and surgeons may be powered by AI or at least assisted by AI.

One of the most significant benefits of AI lies in the **potential of customisation and personalisation**, whereby consumer preferences can be taken into account during the design process, even directly through the customers' voice. This not only is more likely to increase consumer satisfaction, with producers being able to anticipate consumers' needs and preferences based on the data they generate - but it increase customer safety by incorporating the final consumers' personal characteristics and preferences in the design process of the final product through AI, and therefore its intended future use. While its application has so far been rare, an example of AI-powered customisation using voice recognition is the creation of a digital breaker box, where engineers used holographic technology to provide verbal inputs and instructions to design and simulate the prototype based on the preferences expressed by the customer.[118] This is an example of co-creation, where the producer and the consumer work together to design a new product that satisfies the needs of consumers. AI technology however can also simulate how the designed product would operate in the real world and potentially how safe it could be under different situations or uses the consumers make of it.

Developing AI applications such as voice recognition could enhance this level of personalisation even further, allowing for a more direct and personal interaction between AI and individuals, and creating more sophisticated AI-powered personal assistants being able to connect with the other smart devices in a household.[119] These personal assistants could help with various tasks at home, such as operating the other electric devices safely in the house through voice command. Personal assistants may also provide users with information on issues such as energy use, which could help households' optimise energy usage and contribute to climate goals[120]. AI assistants could also help consumers make a better and safer use of their household products.

Finally, another potential benefit of implementing AI in consumer products is that it could be leveraged to protect against cyber-attacks[121] as they are becoming more elaborate and can harm producers' operations to the detriment of consumers. AI may also play a role in detecting, analysing and preventing cyber-attacks affecting production and critical infrastructure.[122] Consequently, Cybersecurity companies have developed AI algorithms to detect viruses and malware and run pattern recognition software.[123]

4.2 Challenges and risks

Although the benefits and opportunities of using AI in consumer products are clear in many cases, there are key questions that need to be answered in relation to the **possible negative implications of AI use on the safety of consumer products.** Although, at present, **much of the debate on this topic is theoretical in nature and evidence of**

[118] Irwin, B. (2018). Mass Customization of Personalized Digital Products.
[119] Emerj. (2020). Everyday Examples of Artificial Intelligence and Machine Learning.
[120] World Economic Forum. (2018). Here's how AI fits into the future of energy.Here's how AI fits into the future of energy.
[121] BCG. (2018). Artificial Intelligence Is a Threat to Cybersecurity. It's Also a Solution.
[122] EC-Council. (2019). Blog: The Role of AI in Cybersecurity.The Role of AI in Cybersecurity.
[123] Laurence, A. (2019). The Impact of Artificial Intelligence on Cyber Security.

real-world examples of these risks being realised is limited, there are a range of challenges and harms that could result from the use of AI in consumer products.

A framework for understanding the possible product safety challenges and risks of AI consumer products is presented in the below figure. In summary, the key characteristics of AI can translate into a range of challenges that have the potential to result in consumer harm of different forms.

This section provides an overview of the relevant characteristics of AI and the potential challenges and harms, before exploring the following topics in greater depth: mutability; robustness and predictability; transparency and explainability; immaterial harms; and the implications for vulnerable consumer groups. These topics were selected in collaboration with OPSS. The methodological approach to this selection of topics is presented in Appendix C.

Figure 4-1: Overview of product safety implications in AI consumer products

4.2.1 Relevant characteristics of AI consumer products

As highlighted in Section 2, AI applications, understood in their broadest sense, share many similar characteristics with other products that incorporate digital technologies (e.g. smart, connected and consumer IoT products). However, particularly considering **ML applications**, there are additional characteristics that are important to note. These main characteristics are described below and the differences between simpler and more complex AI applications are highlighted:

- **Data needs:** AI applications require good quality, unbiased data and a sufficient amount of data to ensure their outputs are robust, accurate, fair and representative. However, the amount of data needed increases as the AI model becomes more complex. As such, more extensive training, validation and testing is needed for ML applications to achieve the same levels of accuracy and predictability in the models deployed as compared with simpler AI systems.

- **Opacity:** For more complex AI applications, there is significant opacity in relation to the visibility and explainability of the algorithms and data used.

- **Mutability:** More complex AI applications enable products to learn and develop over time, instead of relying on explicit instructions. This ability to learn and change the actions and nature of a product inherently requires monitoring and maintenance of AI models after deployment and is a key differentiator between AI and non-AI applications. Although the learning ability of many AI models is reportedly 'frozen'

before placement on the market, updates installed at a later date may effectively change the product and alter the product's risk profile.[124]

- **Autonomy:** Closely linked to mutability, AI applications are increasingly autonomous in their implementation. This ability to make decisions and take actions without human intervention, control or supervision is another key differentiator between AI and non-AI applications.

Building on the above characteristics, the **complexity of AI applications in consumer products is enhanced by a range of factors.** For one, the increasing number of digital components within consumer products and the need for integration or interoperability between products often results in a complex supply chain, with many different economic operators directly or indirectly involved in product development. The ability for frontend and backend control and operation over a product, as well as the use of cloud and edge computing[125], also brings increased complexity when considering AI applications in consumer products.

4.2.2 Potential challenges and harms related to AI consumer products

As illustrated in the above figure, the research suggests a range of potential challenges can occur because of the specific characteristics of (particularly more complex) AI systems. This section presents an overview of the different possible challenges and how they could result in consumer harm, covering challenges related to: robustness and predictability; transparency and explainability; security and resilience; fairness and discrimination; and privacy and data protection.

When considering possible harm, we consider both material and immaterial harm. Material harm refers to physical damage to a consumer and property damage. For instance, this could include an AI-driven robot malfunctioning as a result of automated decisions causing physical injury. Immaterial harm refers to any type of non-physical harm and could include, for instance, replacement of human contact for older people with autonomous products causing mental health issues.[126]

Robustness and predictability. To ensure the safety of a consumer product, it is important for that product to perform as intended by the developer / manufacturer, and as expected by the consumer. However, poor decisions or errors made in the design and development phase can lead to poor algorithmic performance in operation. For instance, these errors can include the selection of inappropriate objectives for an AI system and can translate into issues such as distributional shift, which is where a product fails or struggles to adapt to an environment that is different to its training and testing environment.

These specific issues are examined in greater depth in section 4.3.2, which further explains the concepts related to robustness and predictability and discusses the extent to which and how these issues can impact product safety.

---

[124] European Commission. (2020). Inception Impact Assessment: Proposal for a legal act of the European Parliament and the Council laying down requirements for Artificial Intelligence.

[125] Edge computing allows for local processing and storage of data from edge devices, such as IoT devices, prior to or instead of transmission to the cloud or a data centre, thereby improving latency and reducing costs. See for more information: https://www.networkworld.com/article/3224893/what-is-edge-computing-and-how-it-s-changing-the-network.html

[126] Report from the Commission to the European Parliament, the Council and the European Economic and Social Committee (2020) Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics, COM(2020) 64 final.

Transparency and explainability. As detailed in section 2.2, AI developers are required to make certain decisions, which have a range of trade-offs. One key decision relates to the type of model to employ; simpler, more traditional models tend to be more transparent and interpretable compared to more complex models, but may not be able to achieve the same level of performance. As a result, many AI systems lack transparency and explainability. In addition, a consumer may not know or understand when an AI system is in use and taking decisions or how such decisions are being taken. These issues are discussed in greater depth in section 4.3.2.

Security and resilience. Although the cyber security of AI consumer products is outside of the scope of this study, cyber security vulnerabilities in consumer products have the potential to facilitate or enable consumer harm. For instance, in 2019, European authorities ordered a mass recall of a smartwatch for children via the Rapid Alert System for Non-Food Products (RAPEX). As highlighted in the RAPEX alert, the "mobile application accompanying the smartwatch has unencrypted communications with its backend server"[127], which enabled unauthenticated access to sensitive data, including location history and phone numbers.[128] Although it is unclear whether the smartwatch contained AI, this example illustrates how cyber security vulnerabilities can be exploited to access such devices and their functions (e.g. such as location tracking).

With the development of the consumer IoT market and 5G, the increasing connectivity of, as well as interoperability and integration between, consumer products will also increase the complexity of securing such products. In addition, the risk of losing connectivity could result in safety issues. For instance, if a connected fire alarm loses connectivity, the consumer may not be warned if a fire occurs.[129]

More specifically relevant to AI rather than smart and IoT products more generally, the resilience of an AI system can be tested by other adversarial methods that can present a risk. These methods are characterised by attempts to fool an AI system into mis-classifying certain inputs or making incorrect or inaccurate predictions. A commonly cited example in this respect is in relation to AI tools used by online platforms to identify and moderate illegal or unwanted content posted on those platforms. For instance, researchers have previously demonstrated that removing spaces between words or adding the word 'love' to the end of a word or phrase was enough to trick a system designed to identify hate speech into considering content to be inoffensive.[130] Another study illustrated that neural network powered facial image recognition systems could be fooled into recognising someone as a different person with a high degree of certainty if specially printed multicoloured glasses were used.[131] Although the real-world impact of these kinds of adversarial methods on AI consumer products is unclear, it is a source of possible future risk to AI systems.

---

[127] RAPEX Alert Number: A12/0157/19 – Smart watch for children, 23/01/2019.
[128] CSES and Tech4i2. (2020). Impact Assessment on Increased Protection of Internet-Connected Radio Equipment and Wearable Radio Equipment, Standalone Annex 8: Product-based case studies.
[129] Report from the Commission to the European Parliament, the Council and the European Economic and Social Committee. (2020). Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics, COM(2020) 64 final.
[130] Gröndahl, T. et al. (2018). All you Need is 'Love': Evading Hate-Speech Detection, Proceedings of the 11th ACM Workshop on Artificial Intelligence and Security (AISec) 2018.
[131] Sharif, M. et al. (2016). Accessorize to a Crime: Real and Stealthy Attacks on State -of-the-Art Face Recognition, Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security.

Fairness and discrimination. In many instances, AI systems have been shown to produce discriminatory or inaccurate results, often due to biases or imbalances in the data used to train, validate and test such systems. There are many complexities within this, for instance, the mechanisms for categorisation and classification of data have been shown to contribute to unfair or discriminatory outcomes.[132] Ongoing data collection can also contribute, for instance, by over- and under-representing certain demographic groups, data points or aspects of a phenomenon, as it may not be possible to collect data on or factor all elements of a phenomenon into an AI model. This has been demonstrated in relation to data collected by smartphones.[133] Furthermore, an industry stakeholder interviewed for this study noted that in some instances, pre-trained and tested AI systems are incorporated into products without the manufacturer having a view of the data used and the details of the model.

Although the possible impact of fairness and discrimination issues on physical product safety is unclear, these challenges do have the potential to cause immaterial harms by impacting accessibility and accuracy of outcomes, in particular for vulnerable groups. For instance, individuals with speech impairments have faced issues using voice assistants built into smart speakers.[134] In this respect, consumer representative groups, such as the National Consumer Federation, have highlighted that vulnerable consumers are at greater risk.[135]

Privacy and data protection. The data driven nature of AI systems can lead to privacy and data protection related issues. Inferences made by AI systems have been demonstrated in many environments to seemingly transform non-sensitive data into sensitive personal data and make decisions on that basis. For instance, researchers have demonstrated the ability to use data collected by smartphones to predict clinical depression.[136] Furthermore, linked to the above discussion on fairness and discrimination, AI systems can group individuals based on data received or inferences determined and make decisions based on those categories. Many academics have raised concerns relating to this 'categorical' or 'group' privacy.[137] Furthermore, cyber security vulnerabilities could result in breaches of personal data and related immaterial harm, such as mental health impacts.

This is particularly relevant given that a primary current use of AI in some consumer product sectors, such as utilities, is in relation to the analysis of data collected from consumer devices to inform product improvement and future development. In this regard, industry stakeholders have highlighted a lack of clarity around the legality of transferring certain data, such as operational data. Beyond these issues, however, the relevance of privacy and data protection challenges to product safety is unclear.

4.2.3 In-depth examination of specific AI topics

This section presents the findings on the AI topics selected for in-depth research. These topics build on the above research and cover the following specific issues: mutability;

---

[132] Veale, M. and Binns, R. (2017). Fairer Machine Learning in the Real World: Mitigating Discrimination without Collecting Sensitive Data, 4 Big Data & Society 205395171774353.
[133] Crawford, K. (2013). The Hidden Biases in Big Data, Harvard Business Review. [online]
[134] https://www.voicesummit.ai/blog/how-voice-tech-is-slowly-including-people-with-speech-impediments
[135] National Consumer Federation. (n.d.) Benefit vs Risk in the new digital world.
[136] Farhan, A.A. et al. (2016). Behavior vs. Introspection: Refining Prediction of Clinical Depression via Smartphone Sensing Data, IEEE Wireless Health (WH) (IEEE 2016).
[137] Taylor, L., Floridi, L. and van der Sloot, B. (2017) Group Privacy (eds) (Springer 2017).

robustness and predictability; transparency and explainability; immaterial harm; and product safety implications of AI products for vulnerable consumer groups.

4.2.3.1 Mutability

As highlighted throughout this report, machine learning is a subset of AI that gives computers the ability to learn without relying on explicit instructions. This ability to learn and develop is a key differentiator between AI and non-AI consumer products. This sub-section details how the attribute of mutability is understood in relation to consumer products, before discussing the implications of mutability for product safety and the practices manufacturers take to combat these implications.

What is mutability in a consumer product context

When incorporated into a consumer product, machine learning models can give the product the ability to learn over time and change its actions on the basis of new data. The below box illustrates how this might work through a few examples of machine learning applications in consumer products.

Box 4-1: Examples of mutability in consumer products

For instance, modern **robot vacuum cleaners** contain a range of sensors that collect data about the environment it has been placed in. In fact, a 2019 Review study on Vacuum cleaners[138] developed for the European Commission highlighted at least ten different sensors used by robot vacuum cleaners, including infrared sensors for side and cliff detection, mechanical bumper sensors for collision detection, a tachometer, a gyroscope and an electronic compass. Based on the data collected by these sensors, a robot vacuum cleaner can assess and map its environment, before determining and remembering the most efficient cleaning route; many also contain speech recognition AI systems.[139]

Another distinct example of the use of machine learning in consumer products is the use of deep learning algorithms – a sub-set of machine learning algorithms – by **smartphones to conduct facial recognition**. For instance, in a 2015 paper 'Deep Face Recognition'[140], academics from Oxford University demonstrated the use of a deep convolutional neural network for facial recognition. Industry examples of similar technologies include Apple's Face ID[141] and Facebook's DeepFace[142]. As noted by Apple, such systems, as well as the data they collect, store and use, "will be refined and updated as you use Face ID to improve your experience"[143].

Implications of mutability for product safety

A change to an autonomous action of a product can theoretically result in safety risks as new actions may be learnt and implemented without human oversight, causing physical

[138] European Commission. (2019). Review study on Vacuum cleaners: Final report. Produced by Viegand Maagøe A/S and Van Holsteijn en Kemna B.V. for DG Energy.
[139] Bharadwaj, R. (2019). Artificial Intelligence in Home Robots – Current and Future Use-Cases, Article on Emerj.
[140] Parkhi et al. (2015). Deep Face Recognition.
[141] Apple Computer Vision Machine Learning Team. (2017). An On-device Deep Neural Network for Face Detection.
[142] Taigman, Y., Yang, M., Ranzato, M. and Wolf, L. (2014). DeepFace: Closing the Gap to Human-Level Performance in Face Verification. Facebook Research Publication, Conference on Computer Vision and Pattern Recognition (CVPR).
[143] https://support.apple.com/en-us/HT208108

harm. The European Commission, for instance, has highlighted that this autonomy can "alter a product's characteristics substantially, including its safety features"[144]. However, many of the key sectors highlighted and discussed by the European Commission, including healthcare and autonomous transport, are not within the scope of this study.

In the product groups covered by this study, there is little evidence in the literature reviewed or in the data gathered through interviews of real-world incidences in which the mutability of a product has resulted in physical harm. This raises a few key questions, including: to what extent are models used in the consumer products within scope still learning once placed on the market?; and to what extent do products driven by machine learning change in practice once placed on the market?

Considering the first question, the European Commission noted in its inception impact assessment for its proposed Regulation on AI that "the AI powering a given product would normally not 'learn' or evolve while in operation and instead be 'frozen'"[145]. The comment, noted in a footnote of the report, goes on to state that even if ML capabilities are frozen, software updates may still push ML developments that could alter the risk profile of the product. However, there is limited mention of this same point in the legislative proposal tabled on 21st April 2021[146] or the accompanying impact assessment[147].

Similarly, in the literature reviewed for this study, there is limited discussion on the activities of manufacturers and other economic operators related to this point. The examples of robot vacuum cleaners and facial recognition technology highlighted above suggest that certain models are still learning in operation, but this is not clear. Furthermore, in interviews conducted with industry representatives for this study, it has been noted that OTA updates and upgrades are often used to fix errors or update ML models in consumer products and that, if the changes bring new functionalities or change the risk profile of the product, updated certification documents will be sought.

Therefore, the extent to which ML models that learn in the real world actually change the nature and actions of a product beyond that which is already considered by a manufacturer in development is unclear. To demonstrate, multiple industry stakeholders interviewed for this study highlighted the same point, indicating that they had not, in their products, experienced extensive changes in products post market placement.

However, as described above, there are a range of key decisions that need to be made within the development process that can lead to product safety challenges. These, in some cases, can be exacerbated by the nature of ML models. For instance, the complexity of models that convey mutability often comes with a certain level of opacity. Identifying bugs and other issues in such models can be a costly and difficult process. Furthermore, a lack of transparency in how and why a decision is made or an action is taken can limit the ability to understand the root cause of a defect or harm.

[144] European Commission. (2020). Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics. Report from the Commission to the European Parliament, the Council and the European Economic and Social Committee. COM(2020) 64 final.
[145] European Commission. (2020). Inception Impact Assessment: Proposal for a legal act of the European Parliament and the Council laying down requirements for Artificial Intelligence.
[146] Proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, COM(2021) 206 final, Brussels, 21.4.2021.
[147] European Commission. (2021). Study to Support an Impact Assessment of Regulatory Requirements for Artificial Intelligence in Europe, Final Report (D5)

4.2.3.2 Robustness and predictability

The concept of **robustness and predictability** refers to the ability of an AI application to perform as intended by the developer / manufacturer and as expected by the consumer. This section discusses issues related to robustness and predictability of AI systems and the extent to which and how they can impact product safety. We start by discussing existing AI safety frameworks before drawing on these frameworks to present specific safety issues related to the use of AI systems in consumer products.

In the literature reviewed for this study, similar concepts have been discussed in relation to AI safety. Although much of this research relates to AI safety generally, rather than for products or even consumer products, it raises useful and relevant issues for this study. The below box summarises a selection of these assessments, before one of the most relevant frameworks is presented in more detail.

Box 4-2: Frameworks for AI safety challenges

In 2016, researchers at Google Brain identified four main classes of design errors and five classes of AI safety issues related to the performance of AI systems as intended by the developer.[148] This framework is discussed in more detail below.

In 2018, researchers at DeepMind grouped AI safety considerations according to three categories: i) specification, concerning possible issues related to the purpose of the system; ii) robustness, concerning possible issues related to the ability for the system to withstand perturbations; and iii) assurance, concerning possible issues related to monitoring and control in operation.[149]

In 2019, researchers at Faculty framed AI safety by the relationship between autonomy (considering human controlled vs. autonomous decision making) and intention (considering benign vs. malicious intent). Within this frame, technical and policy problems can be categorised. An example is biased algorithms, which are categorised as a technical problem resulting from benign, human controlled challenges.[150]

As highlighted throughout the above discussions, certain decisions in the development process for AI / ML systems can, if errors are made, result in harmful or unexpected outcomes. In this respect, the Google Brain research is most closely related to the issues we are examining in terms of the AI applications examined. The research identifies four main design and development errors related to: i) the selection of objectives for a system; ii) the method and metrics for evaluating a system's performance; and iii) the use of insufficient or poorly curated training data; and iv) the use of an insufficiently expressive model.[151]

The result of these design and development errors, according to the research, is five possible failure modes. These five classes of AI safety issues include:[152]

- **Negative side effects.** This class considers possible failures related to negative impacts of an ML model on its environment while pursuing its goals.

[148] Amodei, D. et al. (2016). Concrete Problems in AI Safety.
[149] Ortega, M. et al. (2018). Building safe artificial intelligence: specification, robustness, and assurance.
[150] Feige, I. (2019). What is AI safety?
[151] Amodei, D. et al. (2016). 'Concrete Problems in AI Safety'.
[152] Amodei, D. et al. (2016). 'Concrete Problems in AI Safety'.

- **Reward hacking** considers possible failures related to an ML system 'gaming' or taking shortcuts to nominally achieve its objectives.
- **Scalable oversight** considers possible failures related to the need for and ability to provide human oversight to an ML system.
- **Safe exploration** considers possible failures related to how an ML system experiments and learns new actions.
- **Robustness to distributional shift** considers possible failures related to how an ML system acts in an environment different from its training and testing environment.

Implications for product safety

As indicated through the above framework, challenges related to the robustness and predictability of an AI application can occur because of a range of issues in the AI design and development process. More specifically, the need for sufficient good quality data and design decisions, including the choice of objectives and the choice of AI approach, can lead to product safety challenges.

Challenges related to robustness and predictability of an AI application can occur because of the need for a significant amount of data. As the accuracy and relevance of an AI model improves with increased training, validation and testing and with more data, the use of insufficient or inaccurate data may lead the model to draw incorrect conclusions. For instance, one possible consequence in machine learning is overfitting,[153] resulting in an inability to reliably and accurately predict future observations.[154] A practical example could be an AI powered security camera system designed to detect objects and act on that basis (e.g. flag the identification of the object to the user). If the data used to train, validate and test the AI model does not cover sufficient examples of both typical and poorly lit environments, the model may fail to accurately identify objects in failing light.[155]

A further data-related issue that can impact the robustness and predictability of an AI driven consumer product is **bias**. As highlighted in section 4.2, data used to train, validate and test an AI system can be biased in a range of ways. For instance, historical human biases may be inbuilt in a dataset or a dataset may only cover certain demographic groups.[156] This can lead to unexpected and unintended outcomes that are discriminatory or unfair. As an example, researchers have found that certain speech recognition systems face difficulties understanding and responding to users with speech impairments. The issue of the impact of AI systems on vulnerable groups is discussed further below.

From a more practical product safety perspective, biased data could result in distributional shift, as an AI system may be trained to deal with a particular environment based on data that ignores or omits other possible environments that it may encounter. The Google Brain

[153] Overfitting refers to a machine learning model that "models the training data too well", thereby hindering application of the model to new data and negatively impacting a model's ability to generalise. See: Brownlee, J. (2016). Overfitting and Underfitting with Machine Learning Algorithms, Article in Machine Learning Mastery. [online]
[154] Kozyrkov, C. (2020) Training, validation and test phases in AI – explained in a way you'll never forget, Article in Towards Data Science. [online]
[155] Report from the Commission to the European Parliament, the Council and the European Economic and Social Committee. (2020). Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics, COM(2020) 64 final.
[156] Crawford, K. (2013). The Hidden Biases in Big Data, Harvard Business Review. [online]

paper cited above uses the example of a cleaning robot that might learn strategies that are optimised for cleaning one environment (e.g. an office), but might be dangerous when implemented in another (e.g. a factory work floor).[157]

The **design of an AI system** can also undermine the robustness and predictability of an AI consumer product. This can be driven by decisions made in relation to the objectives of an AI system, the evaluation methods or the model choice. As highlighted above, there are significant trade-offs that need to be considered by developers when making these decisions. For instance, simpler models tend to be more interpretable, but may not perform as well (against chosen evaluation metrics) as more complex models. Moreover, the most successful models (in terms of performance) may not be the best option for implementation due to infrastructure considerations, such as the computing power available on a device.[158]

The software engineering notion of 'technical debt', for instance, describes a situation where system designers consciously put off certain actions (e.g. organising and documenting code in an accessible way) in order to reduce time to market or make quick progress. Although the intention is to pay this 'technical debt' back at a later date, it has been argued that technical debt in ML systems can "rapidly accumulate"[159] resulting in hidden and compounding costs related to the challenges of maintaining and improving the system. This is due to the fact that data influences the behaviour of ML models, which can lead to corruption of traditional abstractions and boundaries and debt that is difficult to detect. Furthermore, the inability of typical methods of paying down code (e.g. refactoring code, improving unit tests, reducing dependencies, etc.)[160] are not sufficient in relation to ML algorithms. In practical terms, this can make it difficult to address errors in the ML system over time.

In addition, considering the impact of the design of AI systems, there is much academic debate on the need and utility of more complex AI systems versus simpler approaches. Increasing the complexity and autonomy of AI systems creates additional requirements that can exacerbate challenges related to robustness and predictability. On that basis, some researchers have focused on illustrating the ability to reduce the complexity of systems without reducing utility.[161]

Beyond the impact of 'benign' design and development choices, malicious actions can also result in product safety issues. This can occur in two ways: i) cybersecurity risks, driven by increasing connectivity and complexity of consumer products that use AI; and ii) adversarial methods to fool an AI system into mis-classifying certain inputs or developing inaccurate predictions and outputs.[162]

4.2.3.3 Transparency and explainability

Opacity is often a key characteristic of AI systems and can occur at two levels in relation to consumer products, which raise different challenges for the product safety system: i) opacity towards the consumer regarding the use of AI / ML systems and the

[157] Amodei, D. et al. (2016). 'Concrete Problems in AI Safety'.
[158] Facebook. (2018). Blog: Introducing the Facebook Field Guide to Machine Learning, video series.
[159] Sculley, D. et al. (2015). Hidden Technical Debt in Machine Learning Systems, Google, Inc.
[160] Fowler, M. (1999). Refactoring: improving the design of existing code. Pearson Education India.
[161] See: Ustun, B. and Rudin, C. (2015) Supersparse Linear Integer Models for Optimized Medical Scoring Systems; and Zeng, J., Ustun, B. and Rudin, C. (2017) Interpretable Classification Models f or Recidivism Prediction, 180 Journal of the Royal Statistical Society. Series A, 689.
[162] Algo:aware. (2018). State-of-the-Art Report | Algorithmic decision-making. Report developed for DG CNECT.

autonomous nature of their decision-making; and ii) opacity in the technical approaches used in these AI systems.

A key driver of the first level of opacity is the concept of 'seamless' design.[163] As Mark Weiser wrote, "a good tool is an invisible tool"[164]. When viewed alongside the autonomous nature of such systems, this concept of 'seamless' design poses a challenge to consumers. In the first instance, consumers often do not know whether AI systems are embedded in a device, what data such devices are collecting, whether they are learning and whether they are acting autonomously. Secondly, consumers often do not understand when, why or how a decision has been made autonomously by an AI system. For instance, academic research on the use of consumer IoT products in smart cities has argued that consumers face difficulties knowing if and when products are processing data related to them, or if and when their environment is being altered.[165] This lack of understanding can lead to operational errors by users, as well as unforeseen and unintended use of a product.

Considering the second level of opacity, there are clear trade-offs between the performance and complexity of a model on the one hand, and its ability to produce predictions and outputs that are explainable and interpretable on the other. In this respect, researchers distinguish between two broad categories of models:[166]

- **'Black-box models'** lack interpretability but exhibit greater performance than more interpretable methods. Examples include deep learning methods and ensemble methods;[167]

- **'White-box or glass-box models'** produce explainable results but are not as powerful and struggle to achieve the same levels of performance as more complex methods. Examples include linear or decision-tree based models.

If an AI system works as intended, there is limited concern for its transparency. Challenges occur when something goes wrong. These challenges include the following:

- **Cost of maintenance.** The complexity and opacity of 'black-box' models means that errors and bugs identified in the code are much more difficult to understand, identify and fix than in simpler models. Considering costs, researchers have long understood the relative cost of fixing bugs increases substantially at each stage of the development process;[168] this is only made more difficult and costly when considering opaque AI systems. Similarly, as highlighted earlier in this report, technical debt has been shown to "rapidly accumulate"[169] in ML systems with hidden and compounding costs.

[163] Chalmers, M. and MacColl, I. (1995). Seamful and Seamless Design in Ubiquitous Computing.
[164] Weiser, M. (1994). The World is Not a Desktop, 1 Interactions 7.
[165] Edwards, L. (2016). Privacy, Security and Data Protection in Smart Cities, European Data Protection Law Review.
[166] Linardatos, P., Papastefanopoulos, V., Kotsiantis, S. (2020). Explainable AI: A Review of Machine Learning Interpretability Methods. Entropy 2021, 23, 18.
[167] Ensemble methods use multiple learning algorithms to achieve greater performance than could be obtained solely through any of the constituent learning algorithms.
[168] See for example: Dawson, M., Burrell, D.N., Rahim, E., and Brewster, S. (2010). Integrating Software Assurance into the Software Development Life Cycle (SDLC); or US National Institute for Standards & Technology (NIST). (2002). The Economic Impacts of Inadequate Infrastructure for Software Testing.
[169] Sculley, D. et al. (2015). Hidden Technical Debt in Machine Learning Systems, Google, Inc.

- **Inability to ascertain liability.** If something does go wrong, the lack of transparency and explainability of AI models can impact the ability of the developer of the system, as well as the user, to understand the reasons for an error or malfunction and take action to remedy that error. If physical harm is caused, this has implications for assigning liability. The inability to understand the cause of harm impacts the ability for those that have suffered harm to obtain compensation.[170] Issues of liability, including the impact of opacity, are discussed in more detail in section 6.2.

However, as detailed in the analysis of the market for AI consumer products (section 4), complex AI methods are not commonplace across many of the product groups within the scope of this report. As such, it is anticipated that, at present, the AI being used in most consumer products is simpler, more interpretable, and thus less prone to the challenges highlighted above.

In terms of addressing this opacity, computer scientists and lawyers have for many years been calling for greater transparency[171], leading to significant research focus within academia and industry on improving the interpretability and explainability of complex AI systems. This has resulted in the development of a wide range of interpretability and explainability techniques. For instance, a recent academic publication analysed existing ML interpretability methods and identified four major categories: "methods for explaining complex black-box methods, methods for creating white-box models, methods that promote fairness and restrict the existence of discrimination, and, lastly, methods for analysing the sensitivity of model predictions"[172].

Although a wide variety of methods exist, the authors highlight that the focus of many methods is limited to high-profile AI methods, such as deep learning, concluding that "explainable artificial intelligence still has unexplored aspects and a lot of potential to unlock in the coming years"[173].

In addition, there are examples of the implementation of explainability methods by AI developers. For instance, Amazon noted that a new test called Conditional Demographic Disparity (CDD) influenced its AWS SageMaker Clarify explainability software.[174] Devised by academics from the Oxford Internet Institute, CDD aims to ensure fairness in algorithmic modelling and data driven decisions by mapping the outcomes per demographic group.[175] However, the extent to which such methods are used by developers is unclear.

Beyond industry and academic efforts to improve the transparency and explainability of AI systems, the European Commission's April 2021 proposal for an Artificial Intelligence

[170] Report from the Commission to the European Parliament, the Council and the European Economic and Social Committee. (2020). Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics, COM(2020) 64 final.
[171] See for example: Hildebrandt, M. and Gutwirth, S. (eds). (2008). Profiling the European Citizen: Cross - Disciplinary Perspectives (Springer); and Hildebrandt, M. (2012). The Dawn of a Critical Transparency Right for the Profiling Era, Digital Enlightenment Yearbook 2012, 41.
[172] Linardatos, P., Papastefanopoulos, V., Kotsiantis, S. (2020). Explainable AI: A Review of Machine Learning Interpretability Methods. Entropy 2021, 23, 18.
[173] Linardatos, P., Papastefanopoulos, V., Kotsiantis, S. (2020). Explainable AI: A Review of Machine Learning Interpretability Methods. Entropy 2021, 23, 18.
[174] Zorio, S. (2021). How a paper by three Oxford academics influenced AWS bias and explainability software. Article on Amazon Science: Machine Learning.
[175] Wachter, S., Mittelstadt, B. and Russell, C. (2020). Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI.

Act[176] includes provisions related to transparency. In particular, the following Articles include relevant provisions:

- **Article 13** details requirements related to transparency and the provision of information to users on high-risk AI systems "to ensure that their operation is sufficiently transparent to enable users to interpret the system's output and use it appropriately". Such systems should be accompanied by comprehensible user instructions covering, amongst other elements, the characteristics, capabilities and limitations of performance of an AI system.

- **Article 52** provides transparency obligations for certain AI systems, including emotion recognition systems or biometric categorisation systems. These systems are required to inform users that they are interacting with an AI system.

As these provisions relate to 'high-risk AI systems', they would likely not be relevant to a large proportion of existing consumer products in scope of this study. However, these developments illustrate the growing importance of ensuring transparency and explainability of AI systems.

4.2.3.4 Immaterial harm

In the discourse surrounding consumer AI, some ethicists and interviewees have argued product safety should also cover immaterial, or non-physical, harm. The European Parliament also included a definition of immaterial harm in a recent resolution on a Civil Liability Regime for Artificial Intelligence.[177] It states that "significant immaterial harm should be understood as meaning harm as a result of which the affected person suffers considerable detriment, an objective and demonstrable impairment of his or her personal interests and an economic loss calculated having regard, for example, to annual average figures of past revenues and other relevant circumstances." Other examples of immaterial harm are mental health, data breaches, and threats to one's privacy. Mental health and psychological harms can be perceived as secondary, compared to the more immediate physical harms that a product can cause (such as burns, electrocution, etc.). Financial health is also a significant potential immaterial harm, however it is covered by other regulation, such as the Financial Conduct Authority's regulations. Therefore, this section discusses other types of immaterial risk resulting from AI-driven consumer products and their possible impacts on consumers.

Before expanding further, it is important to clarify, as one interviewee did, that a consumer product such as a smart refrigerator or oven does not collect and process a user's personal data in the same way as a smart TV or smart speaker does. The latter examples require a username and password, often connect to the Internet and a variety of one's different accounts, such as social media and streaming services, thereby increasing the risk of hacking and revealing one's personal information. The former examples are simpler, in that they do collect data that could be triangulated with data from one's other devices, but the risk of a data leak from these devices carries fewer implications for one's personal information; that is to say, the risks differ by product type.

However, there are caveats to this where non-personal data can be used to infer behaviour patterns such as with a smart meter, which collects "energy signatures" that can

[176] European Commission (2021) Proposal for a regulation on laying down harmonised rules on AI (Artificial Intelligence Act) and amending certain union legislative acts COM(2021) 206 final
[177] European Parliament. (2020). Civil liability regime for artificial intelligence (2020/2014(INL)). European Parliament, Texts Adopted.

reveal behaviour and usage patterns[178]. Additionally, hybrid models are increasingly popular, according to manufacturers interviewed for this study. Hybrid models are defined as machines that are run locally, in the user's home, but have certain functions that need to be run remotely on the cloud or in another location. This poses greater privacy concerns, as it can be unclear where one's data is going and what data is being transferred.

This raises the first key category of immaterial harm: **harm to one's privacy and reputation**. When a user's personally identifiable information is not securely stored or transferred, a data breach could lead to this information being accessed, used, deleted or manipulated without consent, which carries several security implications for the individual.[179] Such risks include leaving a user prone to further hacks, such as someone taking control of their social media accounts, manipulating their consumer products, or sharing their address and contact details on a public forum (doxing).[180] In addition, digital libel (false information spread about someone online) or defamation can result from distortion of this leaked data.[181] These are secondary immaterial harms, but should be viewed as connected to the safety one entrusts a consumer AI product to provide.

The second key category of immaterial harm relates to potential **psychological and mental health impacts**, which can occur as a result of attempts to influence consumer behaviour, as well as unforeseen impacts, such as reliance or over-dependence on certain technologies. Recommendation algorithms, for example, have the potential to become prescriptive, as an interviewee noted. Refrigerators or ovens using image recognition technology to understand the types of food someone buys and consumes can learn about a user's eating habits and notify users about whether they are eating healthily. While this appears well-intentioned, and indeed these services are advertised as beneficial for organising a user's consumption habits, it is not the AI's job to prescribe a way of life that, in practice, may not suit their specific user's health needs.[182]

On a societal level, interviewees mentioned another way in which consumer AI can impact a person. The companies implementing these learning capabilities in common household objects use the data these devices gather to digitally "twin" users with products, goods and services they may like based on their data-self. These products are then advertised to consumers on websites, social media, and other platforms they did not realise were connected to how they interacted with their home. This constant monitoring, learning and commodification can be unsettling, and even dissuade people from purchasing these products, as they feel they are being pushed towards certain behaviours and choices for others' benefit.

As interviewees pointed out, there is a more existential risk involved in the incorporation and objectification of these technologies in one's daily life.[183] While relegating human capabilities to a machine does not necessarily pose a safety risk, it does question the role of the human in society if everything is taken care of by machines. An idea often explored in science fiction, these emerging consumer devices are potentially able to demonstrate

[178] Westermann, P. et al. (2020). Unsupervised learning of energy signatures to identify the heating system and building type using smart meter data. Applied Energy. 264.
[179] Cheatham, B., Javanmardian, K., and Samandari, H. (2019). *Confronting the risks of artificial intelligence*. McKinsey Quarterly. [online]
[180] The Cybersmile Foundation. (n.d.). Doxing. The Cybersmile Foundation. [online]
[181] Cheatham, B., Javanmardian, K., and Samandari, H. (2019). *Confronting the risks of artificial intelligence*. McKinsey Quarterly. [online]
[182] Example: Samsung. *Smart Fridge Freezers*. Samsung. [online]
[183] Haddon, L. (2011). *Domestication Analysis, Objects of Study, and the Centrality of Tech in Everyday Life*.

what capabilities, skills, and roles may be lost to automation. This is not always negative, as added convenience may enable a user to spend their time on other tasks and activities while the device handles more cumbersome tasks. However, it is equally important to acknowledge AI's limitations.

What these forms of harm have in common is the negative impact they can have on a user's mental health. Increased distress as a result of using these devices is contrary to the convenience they are intended to provide. As interviewees have pointed out, doing damage control or being notified of repeated immaterial harm wears users down over time. Ofcom research into consumer experiences online found over 80% of respondents had concerns about using the Internet, 37% of those having specific concerns relating to data and privacy.[184] In addition, another study remarked how "people trusted AI algorithms if they were efficient and objective" and ascribed undue authority to these systems, noting that, when something goes wrong, users are likely to experience disillusionment.[185] When embedded in familiar consumer home products, these feelings are likely to be exacerbated when malfunctions, data leaks, or inaccurate recommendations are made.

However, where AI can cause immaterial harm, it can also counteract it. The main example of this is AI-driven privacy-enhancing technologies, which "facilitate data sharing in ways that can improve privacy and in doing so build trust, while personal data stores could help people to exercise more control over their data."[186]

4.2.3.5 Impact on vulnerable groups

In our research on the impacts of AI on the safety of consumer products, a number of groups that may have additional vulnerabilities in relation to AI have been identified. These include: older people, disabled people, people with learning disabilities, and children; as well as those coping with addiction.

When interacting with a new technology product, some members of vulnerable groups may face additional challenges in accessibility or adapting to its role in the home, and the benefits it provides may not be enjoyed by all users. However, many companies have acknowledged this market gap, and addressed it by creating products specifically for vulnerable groups. AI-enabled toys for children, or biometric products for disabled people or people with learning disabilities and older people, are just a few examples of such products; however, for the purposes of this study, we will examine the impacts of AI in products marketed to the general population. This section first discusses the product safety risks and harms that could result from the use of AI-driven consumer products for certain consumer groups, before highlighting the benefits that AI can bring in this regard.

Depending on the nature of the vulnerability—some may be inherent, something one is born with, while others are due to external factors, such as socioeconomic or environmental contexts—there will be different risks involved with using AI-enabled consumer products. The example most often cited by both researchers and interviewees is when these products replace a significant number of activities for the ageing population. From a hair-washing robot to a telecommunication robot, to smart speakers and home assistants that recognise their users' voices, these innovations appear highly convenient

[184] DCMS. (2020). *Policy Paper: National Data Strategy*. DCMS. [online]
[185] You, Y., Kou, Y., Ding, X. and Gui, X. (2021). *The Medical Authority of AI: A Study of AI-Enabled Consumer-facing Health Technology*. Cornell University: Human-Computer Interaction.
[186] DCMS. (2020). *Policy Paper: National Data Strategy*. DCMS. [online]

on the surface.[187] However, they pose two broad risks: reducing user autonomy and increasing reliance on the device. On the former point, one article on care bots mentioned "in-home cameras, facial recognition systems, wearable movement trackers and risk prediction models…undercut the focus on dignity and self-determination central to independent living and community-based care."[188]

On the latter point, as interviewees have commented, these devices are striving toward a validity we normally associate with professional healthcare providers, but they are not able to provide the same level of tailored attention one would receive at a GP or hospital. If users receive positive feedback from a device in their home on their health, based solely on observational data and predictive algorithms, they may not see the worth in travelling to a healthcare provider. Ascribing medical authority to devices and apps could risk misdiagnosis or letting less visible health issues linger and potentially get worse over time.[189] Of course, this is a hypothetical scenario and the exact harms may not be predictable. However, it is still an important concern as they can increase the chances of immaterial harm toward the ageing population.

Another, more dangerous example appeared in a study of an intelligent virtual assistant. When asked a number of health queries on medication or emergency scenarios, these assistants advised users to take actions that could result in harm to their health or death.[190] This raises the greater issue of the reliance of AI systems on historical training data. Historical data can be skewed based on the characteristics of the data subjects at the time, who often represent a narrower portion of society. When a system is trained with these data, it may produce conclusions and recommendations that suit the majority of a given population, rather than individual users.

One case of immaterial harm caused by connected devices and consumer AI that interviewees cited is individuals who are addicted to gambling. A user who expresses an interest in gambling through internet searches, online conversations, and offline interactions with others picked up by a smart speaker or home assistant, may eventually receive recommended advertisements for the lottery or nearby casinos.[191] The device recognises the keywords for gambling and sends advertisements based on this user's profile. This risk can be extrapolated to other people living with addiction, such as alcoholics, who could be bombarded by advertisements for their addiction based on past Internet searches and purchasing behaviours. This algorithmic misunderstanding can severely impact their ability to recover.

Despite these risks, there are a number of benefits these AI consumer products can bring to certain groups. In caring for the ageing population, for instance, a study in which carers were interviewed on their perceptions of AI found that "the three main ways in which robots might be used in eldercare are: a) To assist older adults, and/or their caregivers in daily

[187] van Kemenade, M., Konijn, E. A., & Horn, J. F. (2021). Can we Leave Care to Robots? An Explorative Investigation of Moral Evaluations of Care Professionals Regarding Healthcare Robots. *COJ Robotics & Artificial Intelligence*. 1(3).
[188] Mateescu, A. and Eubanks, V. (2021). 'Care bots' are on the rise and replacing human caregivers. The Guardian. [online]
[189] You, Y., Kou, Y., Ding, X. and Gui, X. (2021). *The Medical Authority of AI: A Study of AI-Enabled Consumer-facing Health Technology*. Cornell University: Human-Computer Interaction.
[190] Bickmore, T.W. et al. (2018). Patient and consumer safety risks when using conversational assistants for medical information: an observational study of Siri, Alexa, and Google Assistant. JMIR.
[191] Busby, M. (2018). Revealed: how bookies use AI to keep gamblers hooked. The Guardian. [online]

tasks; b) To help monitor their behaviour and health; and c) To provide companionship."[192] Interviewees agreed with this stance, mentioning how technologies such as voice and image recognition derive inferences from the data they gather, which may prove useful for monitoring one's health in a connected home. Indeed, previous studies have highlighted how AI-enabled health monitoring apps have improved users' engagement with their treatment plans, and self-management of long-term conditions.[193] These technologies can also benefit users with experience of disability, who may require alternative options to interact with a device; using voice command and recognition allows these users to access and control their devices. However, it is important to maintain a critical eye on prescriptive or instructive AI recommendations, as well as potential privacy concerns arising from storing one's voice and image.

For children, predictive maintenance algorithms and image recognition could prove useful in preventing harmful misuse. One interviewee mentioned how, if a child plays with buttons on a kettle, the device could have a camera on it that recognises the child's age and refuses to turn on. Similarly, using sensors could teach the kettle when it's about to be knocked over, causing it to lock the lid and prevent hot water from spilling out. This could prevent younger users from potential burns or electrocution.

A final safety benefit is, in the case of home security systems, a camera may be able to detect the presence of a fire, and send a notification to users' phones to call emergency services.[194] However, interviewees were quick to add that this requires a level of monitoring and privacy invasion that users may not be comfortable with.

4.2.4 Impact of market trends

Although not explicitly related to AI consumer products, certain market trends are relevant and could exacerbate the above challenges and impact the safety of consumer products. These trends include product re-use and re-cycling and the increasing prominence of online sales channels.

Re-use and recycling. In many digital consumer products, there is a certain built-in obsolescence; after a certain number of years, a manufacturer will stop providing software updates. However, the use of products beyond the end of this support point is increasing, as the markets for re-use, refurbishment and recycling of such products increase and government support for environmentally responsible policies increases. This will result in the presence of insecure and unsafe products on the market. To illustrate the scale of this issue, in 2020, the consumer watchdog Which? estimated that more than 1 billion Android devices globally are vulnerable to attack because they run an Operating System (OS) that has not been supported by security updates throughout 2019.[195] Data from other sources suggest this equates to 42% of Android users worldwide, with more vulnerable populations less well protected.[196]

[192] van Kemenade, M., Konijn, E. A., & Horn, J. F. (2021). Can we Leave Care to Robots? An Explorative Investigation of Moral Evaluations of Care Professionals Regarding Healthcare Robots. *COJ Robotics & Artificial Intelligence*. 1(3).
[193] Stein, N. et al. (2017). A Fully Automated Conversational Artificial Intelligence for Weight Loss: Longitudinal Observational Study Among Overweight and Obese Adults. JMIR Diabetes, Vol 2
[194] Amaryllo. (n.d.). Fire Warning. Amaryllo. [online]
[195] Laughlin, A. (2020). More than one billion Android devices at risk of malware threats. Which?
[196] See: Android.Developers. (2020). Distribution dashboard; and DeviceAtlas. (2019). Blog: Mobile OS versions by country.

E-commerce. According to the European Commission, 69% of EU internet users in 2018 made purchases online.[197] As discussed further in the next section, controlling the adherence to product safety rules in the context of purchases made online is challenging. Key components of this challenge are the presence of new business models, such as online platforms hosting 3rd party sellers, and the increase in the number of consumers purchasing products directly from manufacturers located outside the UK or the EU.[198] Given the focus of these challenges is often cheaper consumer products, AI driven product safety challenges may not be relevant at present. However, over time, as integrating AI into consumer products becomes cheaper, this challenge will become more and more relevant in the context of AI consumer products.

In summary, the characteristics of AI as a technology, including mutability, opacity, data needs, and autonomy, can translate into errors or challenges for the AI system that have the potential to cause harm. These challenges can be categorised according to a range of themes, including robustness and predictability, transparency and explainability, security and resilience, fairness and discrimination and privacy and data protection.

The potential harms can be material or immaterial in nature. Material harms, which are more likely to occur as a result of challenges in the first three themes (i.e. robustness and predictability, transparency and explainability, security and resilience), could include, for instance: an AI-driven robot malfunctioning as a result of automated decisions causing physical injury; AI driven monitoring safety mechanisms in a washing machine failing and causing overheating; or cyber security vulnerabilities in a security vulnerability leading to threats to physical safety. Immaterial harms, which are more likely to occur as a result of fairness and discrimination or privacy / data protection challenges, could include, for instance, replacement of human contact for older people with autonomous products causing mental health issues; or discrimination in access to services for people with disabilities.

However, to date, many of these risks are theoretical in nature and evidence of real-life examples of harm caused by AI consumer products is limited. To some extent, this probably reflects a combination of factors, including: (i) the lack of maturity of many consumer product sectors in using AI; (ii) the due consideration of the possible safety impacts of AI systems by the manufacturers and developers of these products; and (iii) difficulty understanding the role and impact of AI systems when incidences do occur.

Beyond product safety risks specifically linked to AI, general market trends will also bring product safety risks that can exacerbate or be exacerbated by AI consumer products. These include the tensions between built-in obsolescence and the circular economy, and the increasing reliance on e-commerce.

[197] European Commission. (2020). Combined Evaluation Roadmap/Inception Impact Assessment: GPSD and AI.
[198] European Commission. (2020). Combined Evaluation Roadmap/Inception Impact Assessment: GPSD and AI.

5 Regulatory opportunities, gaps and challenges

This section first provides a high-level overview of the regulatory framework for product safety before examining the regulatory challenges resulting from AI-driven consumer products. Following this, section 5.4 presents approaches being implemented globally and by a variety of stakeholder groups to tackle AI risks in consumer products.

5.1 Overview: Regulatory framework for product safety

The UK regulatory framework on product safety is composed of the **General Product Safety Regulations 2005 (GPSR)** and a suite of **product-specific legislation**. The GPSR is a broad umbrella regulation that sets requirements for products to be safe in their normal or foreseeable usage and provides a range of provisions for competent authorities to secure compliance and enforcement when the requirements are not met. The GPSR applies to products intended for or likely to be used, under reasonably foreseeable conditions, by consumers (even if they were not actually intended for them)[199] and if the product is not subject to specific regulations. As illustrated in the below box, certain products, including toys and electrical and radio equipment, are subject to specific regulations. Much of the approach found in the GPSR can be also found in the product-specific regulations, though the latter include additional requirements and obligations.[200]

The main product safety requirements and obligations that can be found in the GPSR and in the product-specific regulations and that are applicable to the argument in this paper are as follows:

- **The safety requirement.** For the GPSR, a "safe product" must not present any risks or only the minimum acceptable risks compatible with the use of the product under a normal or reasonably foreseeable condition of use.[201] The GPSR and the product-specific regulations provide further guidance on the factors that need to be taken into account to determine if a product is safe or unsafe (such as characteristics of the product, instructions, labelling requirements, categories of consumers, or the potential interactions with other products).[202]

- **Producer and distributor obligations.** Producers and distributors must notify in writing to the relevant enforcement authorities if they discover that a product they have supplied is unsafe and poses a risk to consumers. Following the notification,[203] they must take actions to prevent risk to consumers and cooperate with the enforcement authorities which will advise on actions to be taken.

---

[199] The definition of 'product' provided by the GPSR includes products that are supplied or made available to consumers for their own use in the context of providing a service. However, 'product' does not include equipment used by service providers themselves to supply a service to consumers (in particular equipment on which consumers ride or travel which is operated by a service provider).

[200] Masteron, A., Nahon, L. (2018) The UK's consumer product safety legal and regulatory regime. Pinset Mansons, Out-Law Guide

[201] 'Conditions of use' in this context also covers duration of use, putting into service, installation and maintenance.

[202] Masteron, A., Nahon, L. (2018) The UK's consumer product safety legal and regulatory regime. Pinset Mansons, Out-Law Guide

[203] They are also obliged to notify the authorities of the action they have taken to prevent risk to consumers.

- **Obligations for producers (including importers).** The main obligations include ensuring it is a safe product before placing it on the GB market; providing consumers with relevant information to enable them to assess the risk of the product use; allowing traceability; and adopting measures to be informed of the risks of the product. Producers need to make sure that all the potential risks involved in using consumer products are clearly announced in the packaging and instructions provided with the product, including what consumers can do to avoid those risks.

For certain products, additional obligations and requirements are detailed in specific regulations. Where there is a crossover with the GPSR, the product-specific legislation usually takes precedence.[204] However, if the GPSR goes further than the specific regulations in some aspects of product safety then the GPSR will also be applicable to product with specific regulations in that area. The box below provides the list of examples of UK product-specific regulations on product safety that could be applicable to products that can use AI.

Box 5-1: Examples of product-specific regulations applicable to products that can use AI[205]

- **Toys –** Toys (Safety) Regulations 2011 apply to toys manufactured with the following characteristic: toys designed or intended (whether or not exclusively) for use in play by children under 14 years old.
- **Electrical and electronic –** The Electrical Equipment (Safety) Regulations 2016 apply to all electrical equipment that is designed or adapted for use between 50 and 1,000 volts (in the case of alternating current) and 75 and 1,500 volts (in the case of direct current).
- **Machinery –** The Supply of Machinery (Safety) Regulations 2008 ensure that safe machinery is placed on the market or put into service by requiring manufacturers to show how their machinery meet the 'essential health and safety requirements'. For machinery for consumer use the enforcement authorities are the local trading standards authorities or the Secretary of the State while for machinery in use at the workplace it is the Health and Safety Executive[206].
- **Radio equipment –** The Radio Equipment Regulations 2017 require equipment placed on the GB market to comply with a high level of safety (in terms of the health and safety of persons and domestic animals and the protection of property); with an adequate level of electromagnetic compatibility; and to operate in a manner that promotes efficient use of the radio spectrum. For radio equipment for consumer use the enforcing authority is the local Trading Standards authorities.

[204] https://www.gov.uk/guidance/product-safety-advice-for-businesses
[205] The Product Safety and Metrology etc. (Amendment etc.) (EU Exit) Regulations 2019 introduces changes for the Great Britain market. Northern Ireland must align with the EU Directive 2009/48/EC on Toys; EU Directive (2014/35/EU) on electrical equipment; Directive 2006/42/EC on machinery; Directive 2014/53/EU on radio equipment; Directive 2014/68/EU on pressure equipment and assemblies as required by the Protocol.
[206] OPSS (2021) Supply of Machinery (Safety) Regulation 2008: Guidance (GB)

Regulatory mechanisms

There are other regulatory mechanisms in place that support the product safety regulatory framework, for both the products with specific regulations and products under the GPSR. These mechanisms include market surveillance activities, conformity assessment procedures and standards.

Market Surveillance. The UK's Market Surveillance Authorities (MSAs) are both national and local authorities that work in particular sectors and their objective is to protect consumers from non-compliant goods, by ensuring the safety of non-food products placed on the market. Market surveillance is delivered by these MSAs, which are individually responsible for setting risk-based priorities and reporting outcomes for the products pertaining to their sectors. The main MSAs of relevance to consumer products in the UK are:

- **OPSS** is responsible for coordinating market surveillance activity and to oversee the regulatory framework for product safety and acts as a MSA for certain areas of product regulation, such as environmental pollution or energy efficiency (e.g. batteries, eco-design);
- **Local Authority Regulatory Services** for consumer products such as toys, radio equipment, machinery, low voltage electrical equipment etc;[207]
- **Ofcom** is responsible for the regulatory area of electromagnetic compatibility and radio equipment.[208]

Box 5-2: Market surveillance: Challenges and opportunities of AI

AI can play an important role in identifying and detecting unsafe products that are sold online and report them to authorities. The increasing use of e-commerce poses challenges regarding the human resources and capacity of Market Surveillance Authorities to review the products that are placed on the online market. That is why the EU is exploring the use of AI in the form of image recognition technology to protect consumers from dangerous and illegal goods. The technology has been developed by a Danish organisation and is already being used by the Danish Safety Technology Authority. Danish Safety Technology Authority noted that using e-commerce can be difficult for the consumer to see whether a product is legal and safe "therefore, protection is needed to help identifying dangerous goods and to prevent these from being sold. This is where artificial intelligence can create greater consumer safety".[209]

Conformity assessment. For some types of products, manufacturers must follow a process that enables them to make a declaration that the product meets all the requirements that apply to it before they can be placed on the market. Depending on the product, the conformity assessment can either be conducted by the manufacturer or

[207] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/832930/uk-market-surveillance-programme-2019-2020.pdf

[208] https://www.ofcom.org.uk/

[209] KMD (2021) KMD's Image Recognition Technology Aims to Stop Dangerous and illegal protective equipment for sale online in the EU

requires to be conducted by an independent organisation, a notified body, if the relevant regulations applicable to the product provide that.[210]

Standards. One mechanism to demonstrate compliance with relevant product safety regulations is to follow agreed standards for the product's design and manufacture. From January 2021, for goods placed on the GB market, the EU harmonised standards are replaced by the "designated standards" and can be used to demonstrate conformity[211] with GB essential requirements. Designated standards are standards which have been: (i) adopted by any of the following recognised standardisation bodies: the European Committee Standardisation (CEN), the European Committee for Electrotechnical Standardisation (CENELEC), the European Telecommunications Standards Institute (ETSI), and the British Standards Institution (BSI); and (ii) designated by the (BEIS) Secretary of State publishing a reference to them.

5.2 Product safety: Regulatory challenges

AI also presents a range of regulatory challenges related to product safety and liability. These challenges can materialise across all elements of the regulatory regime, including product safety and liability-related legislation, market surveillance regimes, standardisation and accreditation and conformity assessment. This section summarises how the characteristics of AI consumer products described above, as well as more general market trends, can result in regulatory challenges. Liability-related issues are explored in greater depth in section 5.3.

The literature reviewed for this research, as well as the interviews conducted, highlighted a range of **market trends and AI characteristics that could challenge the existing legal texts and regulatory mechanisms on product safety**. These challenges and trends, detailed in section 4, are summarised here:

- **AI-specific challenges.** AI consumer products have certain key characteristics that could lead to product safety challenges and consumer harm. These characteristics (e.g. data needs, opacity, mutability and autonomy), as well as the complexity of the technology ecosystems in which they operate, can lead to regulatory challenges.

- **General trends.** Beyond AI-specific issues, a range of general technology trends can pose challenges to the regulatory framework for product safety. More specifically, these trends include: the blurring of the lines between products and services; the increasing ability for consumer products to cause immaterial as well as material harm; the increasing complexity of the supply chains for consumer products; and issues related to built-in obsolescence and maintenance throughout a product's lifecycle. Although these issues are relevant to AI consumer products, it is important to note that, across the literature reviewed, these challenges are most commonly analysed in relation to software more broadly.

The current regulatory framework for product safety and the existing mechanisms in place to monitor product safety can be applicable to many existing AI consumer products. However, new risks in terms of product safety and liability may be presented with more advanced uses of AI in consumer products. According to an interviewee, at the

[210] From 1 January 2021, UK Notified Bodies automatically become UK Approved Bodies for products placed on the market in Great Britain (in Northern Ireland they are still UK Notified Bodies) and they can carry out conformity assessments to allow the UK Conformity Assessed (UKCA) or UKNI marking to be applied where appropriate for products placed on the UK.
[211] However, the presumption of conformity can be rebutted.

moment, many of the consumer products on the market that use AI are predictable, since they are based on algorithms in a way that for a particular input you will always get a particular output. Therefore, the current regulations can assess the available outputs considering all the inputs given and assess them against the product safety requirements.

In addition, interviewees from all stakeholder groups highlighted the **general suitability of the current regulatory framework**. These stakeholders commented positively on the structure of the framework and supporting mechanisms, as well as the technological neutrality of the framework.

However, gaps in the regulatory regime could arise in relation to AI consumer products that learn and make decisions by themselves once placed on the market. The main challenges and gaps identified through this research relate to: the current definitions and notions detailed in the GPSR and specific product safety legislation; legal uncertainty for businesses; and market surveillance, conformity assessment and standards.

Considering the definitions and notions detailed in the regulatory framework, the characteristics of AI, as well as the general technology trends highlighted above, pose the following challenges:

Product: As illustrated throughout section 4, AI software has the potential to cause a range of safety risks. However, the coverage of software (including AI software) by the GPSR and specific product safety legislation is unclear. For instance, the European Commission's Consumer Safety Network Sub-Group on AI highlighted that the legislation "does not explicitly include or exclude software from its definition"[212].

Although it has been argued that the coverage of software incorporated into a product before placement on the market is clearer, more significant challenges exist related to: (i) the coverage and impact of safety issues resulting from software downloaded on to, or third-party software incorporated into, a consumer product after it has been placed on the market; (ii) the coverage and impact of safety issues resulting from not maintaining software appropriately (i.e. through a lack of updates); and (iii) the coverage and impact of safety issues resulting from changes to a product after it has been placed on the market (e.g. through self-learning AI or software updates).

In addition, the lines between a product and a service are becoming increasingly blurred. More and more products are being sold alongside services, or products are provided to consumers in the context of a service. For instance, remote monitoring, maintenance and updates are ongoing services provided in support of a wide range of products.[213]

These challenges raise questions regarding the clarity of the scope of the regulatory framework and the definition of a 'product', as well as the responsibilities of different entities in the value chain, including software developers.

Safety and harm: According to the GPSR, a product is considered safe when, "under normal or reasonably foreseeable conditions of use […] [it] does not present any risk or only the minimum risks compatible with the product's use, considered to be acceptable and consistent with a high level of protection for the safety and health of persons." Although this definition is considered to be quite broad and could theoretically cover a wide

[212] The Consumer Safety Network's Sub-Group on AI was in fact discussing the EU's General Product Safety Directive (GPSD), which was transposed into UK law as the GPSR. See: Consumer Safety Network, European Commission. (2020). First meeting of the "Sub-Group on Artificial Intelligence (AI), Connected Products and Other New Challenges in Product Safety" to the Consumer Safety Network.
[213] European Commission. (2018). Study on the potential of servitisation and other forms of product-service provision for EU SMEs.

range of consumer harms, product safety has traditionally considered risks to the physical health and safety of the consumer. As detailed further in section 5.3, this perspective is supported by the provisions of the Consumer Protection Act 1987 that relate to product liability, which cover only death and personal injury and physical damage to property. However, as detailed in section 4, AI consumer products have the potential to cause immaterial harms, such as psychological harm.

In addition, given the increasing connectivity of such products, cyber security vulnerabilities can lead to safety risks and this issue is therefore becoming an important consideration. For instance, in 2019, authorities in Iceland identified a children's smart watch that did not pose a direct safety risk, but, as a result of insufficient security measures, could be used to access and potentially cause harm to a child.[214]

Placing products on the market: As highlighted throughout the report, certain AI consumer products are able to learn and evolve once placed on the market. In addition, software updates can change a product and potentially alter a product's risk profile. Although certain provisions of the GPSR, such as Article 20(3), provide some coverage of developing risks that a producer ought to know about, it has been argued that these provisions are not clear in relation to products that use machine learning and software downloads. As such, the legislative focus on ensuring compliance at the point at which a product is placed on the market may not be sufficient in situations where a product has the potential to change autonomously once in the hands of a consumer.

Producer: Given the increasing use of AI and integration of software, the supply chain for consumer products has become more complex. Although the definition of producer in the GPSR is broad, there are potential challenges related to understanding the responsibilities and obligations, as well as liability, of different economic operators in the value chain. For instance, if a user downloads third party software onto a product, which economic operator in the value chain is responsible for ensuring the safety of that combination of product and software? This issue is relevant across both AI-driven and non-AI digital consumer products. Section 5.3 provides more detail on this issue in the context of liability.

These challenges have implications on a range of stakeholders, including businesses, regulatory bodies and consumers. For instance, in interviews conducted for this study, industry stakeholders and regulatory bodies highlighted that, as a result of the above, businesses lack legal certainty on the application of legislation to AI consumer products. In addition, the mechanisms to support businesses in ensuring AI consumer products comply with the legislation are currently lacking. For the most part, product standards do not yet consider the use of AI in such devices and significant challenges exist for conformity assessment bodies and MSAs with regard to AI products. Interviewees from industry noted that these bodies often lack specific knowledge and authority to test and challenge the AI systems being used in consumer products. Moreover, although the growth of e-commerce and online marketplaces is driving growth in consumer product sales, it is also facilitating the prevalence of non-compliant (and unsafe) products on the market.[215]

Beyond these direct challenges to the legislative framework, interviewees from all stakeholder groups highlighted the need to ensure coherence and consistency across UK Government bodies on its approach to regulating AI, as well as smart and IoT products. At

[214] Consumer Safety Network, European Commission. (2020). First meeting of the "Sub-Group on Artificial Intelligence (AI), Connected Products and Other New Challenges in Product Safety" to the Consumer Safety Network.
[215] European Commission, (2020), Combined Evaluation Roadmap/Inception Impact Assessment: GPSD and AI.

present, various UK Government bodies are examining the regulatory issues of AI, as they relate to different issues. Beyond BEIS and OPSS, for instance, related work is being conducted by DCMS, Ofcom and the health service, as well as a range of other public bodies.

5.3 Product liability: Regulatory challenges

At present, consumer product liability issues in the UK are primarily regulated by the 1987 Consumer Protection Act[216], which among others, was the act of Parliament that transposed the European Community (EC) Directive 85/374/EEC on product liability into UK law. These regulations were drafted at a time when commercial product technologies were more straightforward and supply chains less complex, and where the type of damage in focus was mainly on physical harm to life or property. The following section aims to provide an overview of the potential impact that the incorporation of AI into consumer products may present to current product liability rules and determine whether the current product liability regime in the UK is fit for purpose. Or, conversely, if it will need to be revisited in light of the ongoing technological and industrial changes – and see what the options or opportunities for the UK in this respect are.

5.3.1 Impact of AI landscape on liability rules

The evolving changes to products brought by the integration of new technologies, such as Artificial Intelligence (AI), machine learning and robotics, into consumer products are having an impact on the way consumers may experience damage from a product, which in return may impact the attribution of final responsibility.

One of the main technical challenges to the definition of liability rules for products relates to the changing nature of products, which are becoming more complex as they integrate new technologies and actors in the value chain, such as software providers. This exacerbates some of the challenges already contingent to products relying on complex supply chains, as explained by one interviewee. Indeed, when integrated into products, some forms of AI may require changes outside of a producer's control, such as software updates and upgrades, which are needed for the functioning of the products. This is not only a feature of AI, as other products using new technologies depend on future software updates to function appropriately. Moreover, AI and other new technologies present a challenge related to the possibility of products undergoing changes on how they are meant to work after market placement, for example through software updates[217] or alteration through interaction with consumers and their data via machine learning.

In these situations, it is unclear to what extent the manufacturer should be held liable if the damages could not have been predicted.[218] However, in case of third party control over software, as opposed to embedded software which malfunctions and causes damages, it is the third party who could arguably be held liable instead of the producer. This complexity in attributing liability related to foreseeability or control raises questions as to establishing the liable party or establishing a joint-liability regime.

[216] Consumer Protection Act 1987, CHAPTER 43
[217] Expert Group on Liability and New Technologies New Technologies Formation. (2019). Liability for Artificial Intelligence and other Emerging Digital Technologies.
[218] Swanson, G. (2018). Non-Autonomous Artificial Intelligence Programs and Products Liability: How New AI Products Challenge Existing Liability Models and Pose New Financial Burdens. Seattle UL Rev., 42, 1201.

5.3.2 Technical characteristics of AI that pose liability challenges

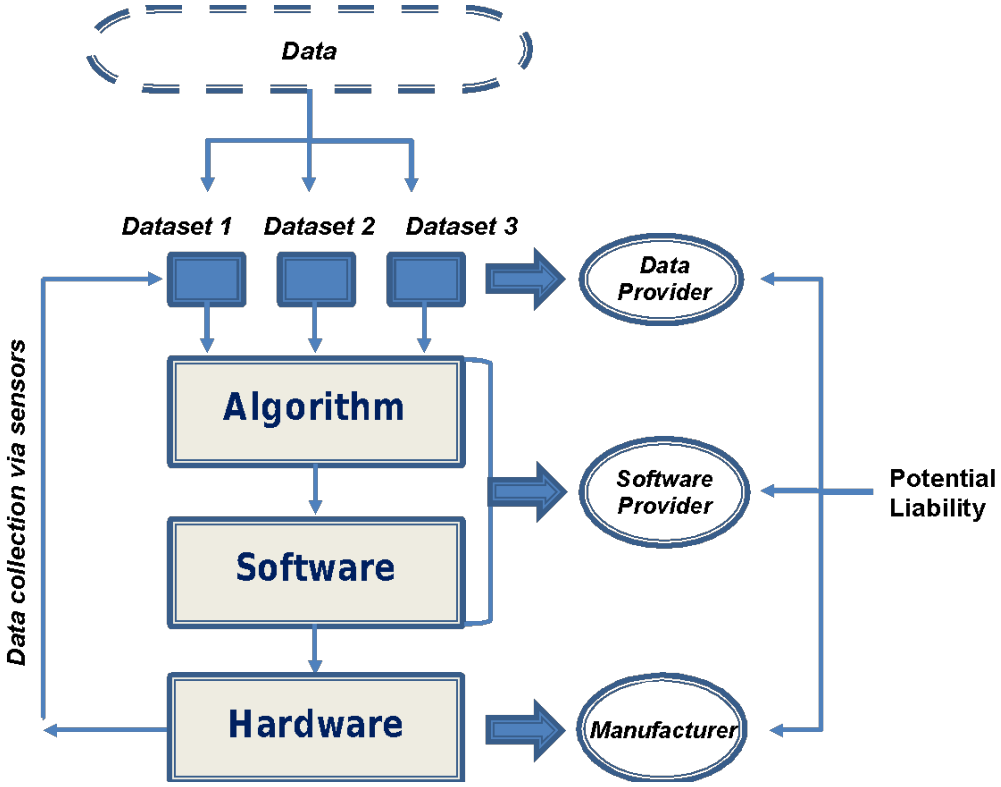
As discussed in previous sections, the added complexity that AI brings into the production processes of consumer products can lead to more actors being responsible for the functioning of consumer products using AI, such as data providers and third party online platforms. This widening web of actors can lead to greater liability issues, because it might be difficult to identify which actor in particular should be considered responsible, and therefore held liable for any malfunctioning of an AI consumer product. Some of the key technical challenges that raise potential liability issues for AI consumer products and AI in general, include the following:

- **Algorithms:** AI may rely on complex algorithms that are opaque[219] and can be difficult to understand by third parties. This makes it challenging to identify the source of potential harm, therefore affecting the attribution of liability;
- **Autonomous systems:** It is possible that products with autonomous systems may change how they operate in unpredictable ways, which may further complicate the attribution of liability. In the case of robotics, for instance, a digital system may harm life or property without human intervention, raising the question of whether a third party should be liable for the decisions of an autonomous system;[220]
- **Openness:** AI products may require frequent updates and rely on external platforms to procure them. This dependency on external providers requires some AI products to be open systems that interact with other actors, which sometimes can be malicious and lead to cybersecurity breaches;[221]
- **Data:** The data-driven nature of AI systems, combined with its testing and training, could result in malfunction, causing physical injury, as a result of using biased, poor quality data, having insufficient data, or poor system design.[222]

These technical challenges contribute to the increasing liability issues stemming from the incorporation of new technologies, such as AI, which make it particularly hard to establish causality between malfunctions in products and the damages they cause. Moreover, this challenge increases as devices and the actors involved in their operation become increasingly interconnected.[223] To illustrate this complexity, the following diagram shows the different stages in the value chain where liability could be attributed in the case of AI product malfunction, namely: (i) at the data level; (ii) at the algorithm and software level; and (iii) at the product level.

[219] Center, E. (n.d.). Epic - ai and human rights. Retrieved April 29, 2021, from https://epic.org/ai/
[220] Zech, H. (20210. Liability for AI: public policy considerations. ERA Forum 22, 147–158 (2021). https://doi.org/10.1007/s12027-020-00648-0
[221] Report from the Expert Group on Liability and New Technologies – New Technologies Formation (2019)
[222] CSES, (2020), Framing the Nature and Scale of Cyber Security Vulnerabilities within the Current consumer Internet of Things (IoT) Landscape, study for DCMS – included a case study on ICT-facilitated abuse.Framing the Nature and Scale of Cyber Security Vulnerabilities within the Current consumer Internet of Things (IoT) Landscape
[223] Zech, H. (20210. Liability for AI: public policy considerations. ERA Forum 22, 147–158 (2021). https://doi.org/10.1007/s12027-020-00648-0

Figure 5-1: Distribution of potential liability risks in the AI value chain



The above diagram shows the different stages at which an AI system might malfunction and where potentially different actors could be considered liable. For example, while traditionally manufacturers are considered the primary liable party in the event of damage, this is complicated in the context of third party suppliers whereby damages might be due to the malfunctioning of software which may fall under the responsibility of third party suppliers and are out of the control of the final manufacturer. However, software relies on the operation of algorithms which in turn rely on data to operate.

Consequently, the source and type of data that algorithms use may also impact liability questions, as the quality of an algorithm's performance may only be as good as the data it processes. This factor might potentially affect the attribution of liability to third-party software developers and providers. For example, in cases where the software provider is not responsible for the data and quality that feeds into its algorithm, such as in a self-learning algorithm adapting to consumers' lifestyle data, it might prove difficult to attribute full liability to the software provider given that its algorithm might be functioning correctly, but it is the consumer who may be providing dangerous data through risky behaviour.

One of the stakeholders we interviewed provided an example illustrating these risks in the following terms: 'what happens when an algorithm behind a smart heating controller malfunctions and overheats a house, even though it nominally is basing its decisions on the customer's behaviour?' The potential impact of misuse by the consumer may put attention on the consumer's own use of a product and whether a plaintiff did everything in its ability to avoid injury. Indeed, manufacturers might have to prove misuse from users to show that its products' design is not defective. This will likely give greater salience to the importance of being transparent when placing products that incorporate AI, by presenting

the dangers associated with the use of its AI products (both pre- and post-sale) as well as any foreseeable misuse of the product.[224]

In the event of a malfunction that could be due to product misuse, it would be particularly challenging for an injured party to prove a product malfunction, due to the lack of explainability for the malfunction, given that the product was designed to model its behaviour based on the injured party's behaviour. This is particularly a risk in the case of deep learning types of AI, where a 'back box' algorithm would potentially leverage complex deep neural networks and significant amounts of data as inputs to produce an output, through processes outside the purview of third parties involved and is difficult even for data scientists, programmers and especially users to interpret given its complexity.[225] In other circumstances, such 'black box' algorithms may be proprietary and kept hidden in order to protect intellectual property and retain a competitive advantage. Examples of these 'black box' proprietary algorithms include Google Search's ranking algorithm, or Amazon's recommendation system.

However, in a liability regime, where the burden of proof lies with the victim to prove malfunction and damage – the fact that there are complex black-box AI systems that are not explainable or accessible to data scientists, let alone lay consumers, might prove a challenge in not only attributing liability but also in claiming compensation as the responsible parties cannot be determined. Indeed, it may even be difficult to establish whether any malfunction occurred at all with a particular part or component, as victims may not be able to trace the defects in the algorithm that caused the damage. This added complexity makes it more difficult for the injured party to satisfy the legal conditions and discharge the burden of proof to establish liability and gain compensation for damage. This has led to calls from some stakeholders for the reversal of the burden of proof in instances where it is too complex and onerous to expect injured parties to prove damages are due to the product malfunctioning.[226]

5.3.3 Revision of liability concepts and definitions in the context of AI

The issue of liability in the context of AI is consequential to all operators in the supply chain of an AI-driven product, in case unclear liability rules may impart liabilities unfairly, if not wrongfully. A lack of clarity on liability issues might also affect future trust in AI consumer products, as it may seem too risky to incorporate AI functions. Ineffective liability rules for AI products also present risks that may dent the confidence of consumers as any issues or complexities in settling liability issues between different manufacturers, suppliers, and third party providers may delay the payment of compensation claims, or discourage consumers seeking legal redress.

Many of the potential challenges related to the attribution of liability lie in definitional and conceptual issues that have been defined in past product legislation such as the GPSR[227] or the EU's Product Liability Directive[228], which was implemented by the 1987 Consumer Act.[229]These may need to be re-evaluated to develop a liability framework that is fit-for-

[224] BLG. (2021). Artificial Intelligence and Product Liability: Catching up with the Future
[225] Dickson, B. (2020). The dangers of trusting black-box machine learning
[226] Expert Group on Liability and New Technologies New Technologies Formation. (2019). Liability for Artificial Intelligence and other Emerging Digital Technologies
[227] General Product Safety Regulations 2005
[228] Council Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products
[229] Consumer Protection Act 1987

purpose to consumer products incorporating AI. Examples of terms and concepts that might be affected by the incorporation of AI may include, but not be limited to the following:

- The concept of '**product**' itself should be considered in order to determine whether a product may be regarded as complete when it is sold without software that is downloaded or incorporated at a subsequent stage. Indeed, it should be clarified whether software itself should be considered a product regardless if it is embedded in a tangible product prior to or after its placement on the market. An interviewee described it in the following way: '[we] don't have a fixed understanding of what a product is. It's more of a continuous development'.

- Another concept that might need to be reconsidered is that of '**producer**', in terms of who might be the party that may be considered liable for an AI product which malfunctions – especially if the AI responsible is provided externally. Another challenge would be posed by 3D printing, where in case of product malfunction it might prove challenging to identify the producer; whether this should be the provider of the original product designs, the producer of the printer or an amateur user of a 3D printer, in non-commercial settings.[230]

- The definition of '**damage**', as well as '**defect**', may also need further consideration, as notions of harm may increasingly include risks that have 'non-physical' effects such as damage to personal data.[231]However, it may be difficult to disentangle the liability aspect of damage to data from data protection regimes.

One of the interviewees in this study made the following assessment concerning the varied terminological issues by stating that every 'moving part' should be considered a product and that in the case of AI, there are three moving parts: hardware, software and learning software, which should be factored more clearly in the current liability regime. In light of the definitional issues related to AI technologies and the liability questions they raise for consumer products, it would be helpful to assess whether the current liability regime covering consumer appliances in the UK is fit for purpose or whether it requires adaptations to take into account the changes brought by new technologies? A question that needs to be investigated is whether a separate safety framework should be elaborated specifically for AI products, distinct from non-AI products?

5.3.4 Impact of AI on the liability framework for consumer products in the UK

Our consultations with stakeholders suggest that the UK's product liability regime affords adequate protection to consumers generally. However, it may be the case that that the current liability rules need to be reviewed and clarified for products incorporating AI, according to a few of the stakeholders interviewed. One interviewee, for instance, suggested that there may be some ambiguity in the current rules on whether a product incorporating AI falls under a different liability regime as opposed to a version of the same product which does not incorporate AI. This ambiguity may apply to any changes to a product made through subsequent software downloads or updates, with one interviewee from a UK public body suggesting that 'anything that is downloaded' provides a challenge to the current common law interpretation of a product and the applicability of existing product liability rules.

Uncertainty about the applicability of product liability laws arises partly due to a lack of clarity about whether software constitutes a product or a service, which may require

[230] Austin, A. (2020). Product liability in the AI age.
[231] BEUC. (2020). Product Liability 2.0 - How to make EU rules fit for consumers in the digital age

changes to the rules.[232] Consequently, should a claimant in the UK experience damage from a product powered by AI, there is a risk that he or she may encounter complications in court when claiming compensations under the current liability regime if, for example, the damage was caused by a product relying on third party software. This potential ambiguity present the risks that users of standard non-AI household appliances would be protected differently from those using similar AI-powered household items relying on third party software. However, it is difficult to establish the extent to which these ambiguities currently present a real threat to UK consumers as there have been few reported examples of liability issues in products with AI being treated by UK courts.[233]

One interviewed stakeholder suggested that the impact of this is that producers might be discouraged from introducing AI in their current products for fear that the lack of legal certainty concerning liability rules might expose them to legal challenges, with another Government stakeholder highlighting: ‘you don't want someone considering whether to put AI in a device or not based on liability’. There is therefore a concern that unclear liability rules for products incorporating AI might discourage manufacturers and producers from entering the AI market, and impact companies’ willingness to invest in AI consequently also affecting innovation.[234]

However, despite the fact that AI features pose new challenges and raise questions about whether all aspects of the existing liability regime are sufficient in dealing with damage caused by AI in consumer products, as mentioned previously, the current framework in the UK is considered reasonably comprehensive for non-AI related liability issues by the stakeholders we spoke to on the subject. Moreover, interviewees discouraged the suggestions of creating a separate set of liability rules for AI products.

One of the critical characteristics of the current regime that seems to work well for non-AI consumer products is the application of 'strict liability' to liability claims. In a regime based on strict liability, the liability falls on the producer even if it acted in good faith unless the cause of the damage was not predictable at the time a product was placed on the market. This means that consumers are protected from all foreseeable risks at the time, but that producers may benefit from protections should technological development evolve in ways that could not have been perceived before a product was introduced. The predictability requirement in the current liability regime may, however, be challenged by new technologies such as autonomous devices or the integration of learning capabilities. This is because the creation, however far-off, of machine learning that can learn and take decisions on its own raises the issue of its own responsibility and therefore legal personality. This leads to the addition of a potentially new liable party in the AI value chain, the AI itself, which is outside of the control of its creators and which therefore adds to challenges faced by regulators.[235]

One interviewee, a UK public body, recommended the current system in the UK to be updated to better cover AI products as well in the future, as this would extend the comprehensive liability regime which is present for traditional consumer appliances to the new appliances integrating AI that are entering the market, instead of creating new liability rules to deal with AI-powered products. The same interviewee argued that the same strict liability should be maintained, as any stricter requirement or absolute liability might stifle innovation.

[232] Austin, A. (n.d.). Product liability in the AI age.
[233] ICGL. (2020). England & Wales Product Liability Laws and Regulations
[234] Pinsent Masons (2020). Overcoming barriers to AI adoption – liability
[235] Austin, A. (n.d.). Product liability in the AI age.

Extending the current liability framework to products with AI would imply that the rules on the burden of proof would be maintained, where the responsibility of proving that damages inflicted were due to defects in a product falls on the consumer. The alternative, a 'reversal of the burden of proof', as mentioned in 5.3.2, would require producers having to prove that their products are not defective, and would therefore possibly present a disadvantage by putting excessive burdens on companies, in particular SMEs. One interviewee explained that reversing the burden of proof might tilt the liability regime towards favouring consumers, even though producers are more likely to be able to defend themselves and their products due to their technical knowledge.

Other issues that need to be investigated include the current liability exemptions and whether these should be maintained for AI products, such as the 'development risk defence' which exempts the producer from liability in cases where the state of scientific and technical knowledge at the time when the product was put into circulation was not such as to enable the existence of the defect to be discovered. Keeping this exemption would continue to offer significant protections to producers of AI products.

5.3.5 Evolution of future liability issues in AI consumer products

It is unclear yet what liability problems may arise in the future, as this is still a very new area that has not been looked at by regulators in the UK and elsewhere in great depth as yet, so there is little evidence to date of which characteristics of AI are likely to be more problematic if the safety framework for consumer products is not adapted. It can be foreseen that as products using AI proliferate, the demand for clearer liability rules will become stronger as AI technology matures and its risk become more clearly understood. Currently, there are few examples in the world of geographies trying to regulate liability issues with AI. Steps have been taken in the UK to assign liability rules for driverless cars[236], the same has not yet happened for any other consumer products. However, some jurisdictions such as the EU are currently carrying out consultations aimed at determining the liability risks posed by the incorporation of AI into products and exploring the possibility of updating the current liability rules. In addition, New Zealand[237] and Australia[238] have extended their liability rules to cover software and treat them as 'goods'.

An area of interest that may affect the future legal remedies that are sought could relate to claims being made due to discrimination.[239] This is particularly the case if perhaps future AI relies on biased data, affecting its behaviour from individual to individual. The societal implications of such developments within AI are expected to be closely monitored by policymakers and civil society over the coming years.

5.4 Approaches to tackling AI risks in consumer products

The challenges and risks of the use of AI in the safety of consumer products raise the question of how the current regulatory framework should be modified to tackle these issues. Although legislation tends to be technology-neutral to avoid the obsolescence of their rules, there have been some developments at the national, European, and international levels to address these challenges. This section provides an overview of these developments.

[236] Automated and Electric Vehicles Act 2018, CHAPTER 18
[237] Consumer Guarantees Act 1993
[238] Dundas Lawyers. (2017). Software licences held to be "goods" under ACL, available at: https://www.dundaslawyers.com.au/software-licences-held-to-be-goods-under-acl/
[239] Scott-Lawler, H. (2020). Artificial Intelligence ('AI'): Legal Liability Implications.

5.4.1 Standardisation

Standards can be a tool for self-regulation by the industry and other key stakeholders, who can define for themselves the requirements for products, laying down technical procurement and product development. But standardisation can also serve as a tool for the policy maker. Legislation can refer to standards bodies, such as the ISO, CEN/CENELEC or national standards bodies, to bring solutions to the market while implementing new ideas and innovations to products. Standards allow transparency and trust in the application of technologies, and at the same time support communication between all parties involved by using uniform terms and concepts.[240]

Standards are the result of national, European or international standardisation work and are developed by committees according to defined principles, procedures and rules of presentation. All stakeholders, such as consumers, manufacturers, businesses, universities, research institutes, authorities, testing institutes, etc., can participate in the work of these committees and anyone can contribute ideas during the public consultation stage of the standards development process. The experts within the standards committees develop standards contents by consensus and taking into consideration the stakeholders' interests. The robust process provides a high degree of legitimacy for the standards. Once published, standards are subject to regular review so can be updated as technology develops or withdrawn if they are no longer appropriate.

Standards, therefore, provide the institutional infrastructure needed to develop new technologies and safety procedures in a controlled manner, including the research and development of AI.[241] Since they can promote rapid transfer of technologies from research to application, and to open international markets for businesses and their innovations, standards can play a key role in creating a framework for AI.[242] Thus, standards can serve as a tool for policy makers who could set the objectives and use standards to develop state-of-the art measures to achieve those objectives.[243] Standards might also be used to accelerate the development of technology and allow policy makers to consider where legislation might be required. Standards are often not legally mandatory but can be made mandatory by law, therefore, there is a role of standards as a precursor to regulation.

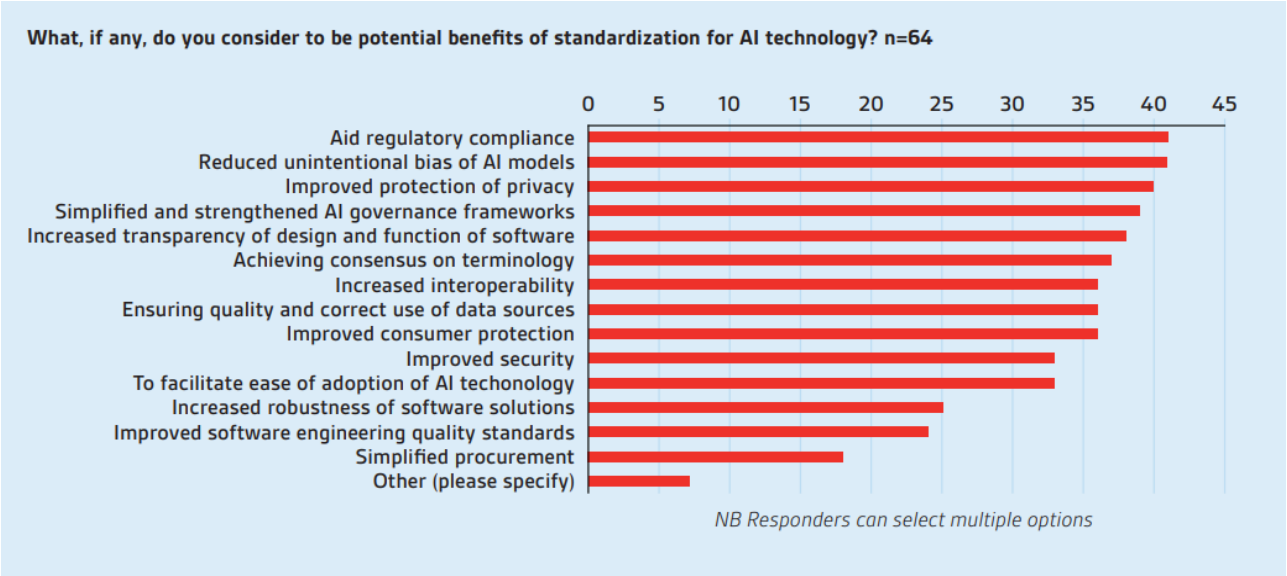
The relevant recent developments or work on standards for AI products is being carried out at national, European and international levels.

At the **national level, the UK** published in 2017 an independent review on how AI industry can be grown in the UK, in which it explicitly mentioned the need for standards that provide guidance on how to explain decisions and processes enabled by AI.[244] The same year, the first committee dedicated to AI was formed by BSI.[245] The committee, which mirrors the work of ISO/JTC1/SC 42 brings together stakeholders to agree on best practice and develop standards for the industry. The focus of the committee is moving from broad technical questions into topics such as governance and bias,[246] as these were the issues

[240] DIN DKE (2020) German Standardization Roadmap on Artificial Intelligence.
[241] Cihon, P. (2019) Standards for AI Governance: International Standards to Enable Global Coordination in AI Research & Development.
[242] DIN DKE (2020) German Standardization Roadmap on Artificial Intelligence.
[243] Cihon, P. (2019) Standards for AI Governance: International Standards to Enable Global Coordination in AI Research & Development.
[244] Hall, D.,Presenti, J. (2017) Growing the Artificial Intelligence Industry in the UK
[245] https://standardsdevelopment.bsigroup.com/committees/50281655
[246] BSI (2019) BSI's activities on Artificial Intelligence (AI)

perceived as most relevant to standardisation for AI technology in a survey conducted by BSI.

Figure 5-1: Benefits of standardisation for AI technology, BSI survey data



Source: *BSI White Paper – Overview of standardization landscape in artificial intelligence*

In relation to specific standards activities, in 2016 the first British Standard on ethical design of robots was published.[247] This provides guidelines on how to identify potential ethical harm in the design and application of robots and autonomous systems.[248] For the healthcare sector, BSI and the MHRA developed recommendations for how standards can support regulatory frameworks and the need for guidance on AI solutions for patient care.[249] Currently, BSI together with BEIS, the Better Regulation Executive (BRE) and the UK Quality Infrastructure (UKQI) are involved in the development of "Standards for the Fourth Industrial Revolution Action Plan" the aims of which are that regulations and standards ensure innovation while also meeting and being adapted to government plans and industry opportunity[250]. However, the extent to which this would be applicable to AI consumer products is unknown at the moment.

Another work of note comes from the Alan Turing Institute and the Centre for Data Ethics and Innovation, which reflect the UK's commitment to shape standards that could govern the ethical performance of AI products.

At the **European level**, the CEN/CENELEC Focus Group on Artificial Intelligence is a body that was created to coordinate activities and to identify the need for standards where not already covered by ISO/IEC SC42. It was established in 2019 by CEN and CENELEC as a temporary working group with the task of developing a roadmap for AI standardization at European level. The Focus Group does not develop standards but identifies specific European requirements for AI, such as terminology, ethics and safety and sector specific standards.[251] [252] The Group was created following EU's strategies such as the European

[247] BS 8611
[248] BSI (2019) BSI's activities on Artificial Intelligence (AI)
[249] BSI MHRA AAMI (2019) The emergence of artificial intelligence and machine learning algorithms in healthcare: Recommendations to support governance and regulation
[250] CQI (2021) Innovation, Standards and the battle for global competitiveness
[251] CEN-CENELEC (2020) Road Map Report on AI, version 2020-09.
[252] CEN-CENELEC (2020) Response to the EC White Paper on AI, version 2020-06.

Commission Rolling Plan on ICT standardisation, which recommended fostering coordination of standardization efforts on AI in Europe, promoting interaction of all stakeholders taking into account their vision and real needs while ensuring coordination between standardisation efforts on AI in Europe and other international standardisation efforts.

At the **international level**, IEC and ISO have already set up a joint committee, ISO/IEC JTC 1/SC 42 "Artificial Intelligence", which is the central body for AI standardisation and is responsible for the development and publication of international standards on AI. The joint committee is developing standards on AI. For instance, the joint committee has already published "Overview of Trustworthiness in Artificial Intelligence (ISO/IEC TR 24028)" and is developing standards such as "Functional safety and AI systems (ISO/IEC AWI TR 5469)" or "Objectives and approaches for explainability of ML models and AI systems (ISO/IEC AWI TS 6254)". However, the extent to which these are specific to consumer products is unclear.

Besides SC 42, other ISO and IEC committees also have ongoing standardisation activities for AI, such as ISO/IEC JTC1 SC 27 (Information security, cybersecurity and privacy protection), ISO/IEC JTC1 SC 7 (Software and systems engineering), ISO TC 199 (Safety of machinery), ISO TC 22 (Road vehicles), ISO TC 215 (Health informatics), IEC TC 65 (Industrial-process measurement, control and automation) and IEC SEG 10 (Ethics in autonomous and artificial intelligence applications).[253]

5.4.2 European and international regulatory developments

Some leading countries in AI, such as US and China, are closely looking at the standardisation activities to take steps towards regulating AI in the context of product safety. This section presents a summary of the main regulatory developments in the European Union, United States and China.

European Union

The EU has taken a range of steps towards regulating AI generally as well as in the context of product safety. The European Commission recently published the Proposal for a Regulation laying down harmonised rules on AI (Artificial Intelligence Act) and amending certain Union legislative acts.[254] The proposal, published on the 21st of April 2021, lays down rules on providers that place AI systems on the market (including providers established outside the EU) and it also applies to users of AI systems in the EU or that are located in a third country where the output produced by the system is used in the EU. Therefore, the proposal affects all parties involved in product safety: providers, importers, distributors, and users. The proposed regulation would complement other horizontal EU legislation such as the General Data Protection Regulation (GDPR), for the privacy issues that AI systems may cause, and sectoral safety legislation, like machinery products (the current Machinery Directive will be updated by a Regulation for Machinery Products in which it will classify as 'high risk' the AI systems that are used as safety components).

The EU aims to overcome some of the challenges that AI may pose to safety, European values, and fundamental rights and principles. The proposed regulation introduced a four-tiered risk framework that classifies the risk posed by AI systems to the health, safety, and

[253] CEN-CENELEC response to the EC White Paper on AI, version 2020-06 https://www.cencenelec.eu/news/policy_opinions/PolicyOpinions/CEN-CLC%20Response%20to%20EC%20White%20Paper%20on%20AI.pdf

[254] European Commission (2021) Proposal for a regulation on laying down harmonised rules on AI (Artificial Intelligence Act) and amending certain union legislative acts COM(2021) 206 final

fundamental rights of users. Therefore, according to this framework AI systems could pose 'minimal or no risk', 'limited risk', 'high risk', and 'unacceptable risk'. The AI systems that pose a high risk will be required to undergo conformity assessment before and after being placed on the market. For AI systems that pose unacceptable risks, the proposal lists four types of AI practices that are prohibited. Some of these practices, particularly 1 and 2, could be relevant to consumer products[255]:

1. Placing on the market, putting into service or use of an AI system that deploys subliminal techniques beyond a person's consciousness in order to materially distort a person's behaviour in a manner that causes or is likely to cause that person or another person physical or psychological harm.

2. Placing on the market, putting into service or use of an AI system that exploits any of the vulnerabilities of a specific group of persons due to their age, physical or mental disability, in order to materially distort the behaviour of a person pertaining to that group in a manner that causes or is likely to cause that person or another person physical or psychological harm.

3. Placing on the market, putting into service or use of an AI system by public authorities or on their behalf for the evaluation or classification of the trustworthiness of natural persons with the social score leading to detrimental or unfavourable treatment that is either unrelated to the contexts in which the data was originally generated or unjustified or disproportionate.

4. Use of "real-time" remote biometric identification systems in publicly accessible spaces for law enforcement purposes, subject however to broad exemptions that, in turn, are subject to additional requirements, including prior authorization for each individual use to be granted by a judicial authority or an independent administrative body in the member state where the system is used.

In addition, there are provisions relating to transparency to ensure people know they are dealing with an AI system[256] but also enable users to interpret the system's output and use it appropriately[257].

This new legal framework on AI is combined with a new coordinated plan with EU Member States[258] that outlines the necessary policy changes and investment at Member States level to strengthen Europe's leading position in the development of human-centric, sustainable, secure, inclusive and trustworthy AI. Both measures build on the Commission's White Paper on AI [259] published in 2020, in which the EU's vision for AI was set out.

United States

The US also has a leading position in AI. In 2016, the US presented its first national AI strategy, as well as launching a $2 billion funding initiative, called "AI Next", for a period of five years to develop the foundations for the next generation of AI systems by the Defense Advanced Research Project Agency of the US (DARPA). DARPA aims to promote a new wave of AI systems that are more robust and trustworthy than previous systems and to overcome AI development, which the US government believes has, in many countries,

[255] Article 5,m .
[256] Article 52
[257] Article 13
[258] COM(2021) 205 final
[259] White Paper on AI

recently focused too much on machine learning and to develop a new generation of autonomous systems that can also work in teams with humans.

In relation to the safety of AI products, the Consumer Product Safety Commission (CPSC) with the current administration, aims to shift away from the deregulation approach of the previous government and move towards increased scrutiny and enforcement of consumer products. The CPSC is gathering information from manufacturers and importers of consumer products that use AI, and other stakeholders, on how to regulate current and potential safety concerns in AI products. For that purpose, CPSC has developed the following list of considerations[260]:

- **Identification**: Determine presence of AI and machine learning in consumer products. Does the product have AI and machine learning components?
- **Implications**: Differentiate what AI and machine learning functionality exists. What are the AI and machine learning capabilities?
- **Impact**: Discern how AI and machine learning dependencies affect consumers. Do AI and machine learning affect consumer product safety?
- **Iteration**: Distinguish when AI and machine learning evolve and how this transformation changes outcomes. When do products evolve/transform, and do the evolutions/transformations affect product safety?[261]

CPSC is taking into account existing safety standards for AI, including Underwriters Laboratories (UL) and the International Organization for Standardization (ISO) and might also look at EU's recent activity on AI.[262]

China

China plans to become the leading AI nation in the world by 2030 and is setting economic targets for this move. AI development is heavily dependent on the availability of large volumes of data, and in this respect, China's data protection framework supports a state-led push for innovation and growth. In relation to standardisation, observers have noted that China is pushing for its own distinct set of national standards in areas such as the internet of things and cloud computing.[263] In 2018 China published the White Paper on AI standardisation in which it highlighted the importance of standards for AI products. It also raised specific concerns regarding:

- Safety and security, especially in relation to algorithmic harm and the difficulty of controlling algorithmic development;
- Ethics – The paper states three agreed principles: human interests, liability and consistency of rights and responsibilities;
- Privacy issues – The White Paper noted the need of a clear definition on privacy issues.

[260] Swanholt, E., McGaver,K. (2021) *Consumer Product Companies Beware! CPSC Expected to Ramp up Enforcement of Product Safety Regulations*
[261] National Law Review (2021) The CPSC Digs in on Artificial Intelligence
[262] National Law Review (2021) The CPSC Digs in on Artificial Intelligence
[263] Hogan Lovells (2018) AI and your business: A guide for navigating the legal, policy, commercial and strategic challenges ahead.

In contrast to the EU, the approach followed by other countries, such as Japan, is less focused on developing binding AI specific legislation and more on producing guidelines that address key principles while also being consultive with industry. The Japanese Government's guidelines were one of the bases of the OECD's principles and have focused on education on the issues and challenges of AI.[264]

Countries might have been hesitant to introduce AI-related regulations and one of the reasons might be the fear of obstructing innovation. On the other hand, new regulations have been influential across countries, as has been the case of the EU GDPR, which has influenced regulation in other continents. Therefore, we might expect regulatory developments to happen sequentially in various countries, at least to the extent that their governments share similar values.

5.4.3 Industry and non-legislative approaches to tackling AI challenges

Apart from formal standardisation, there are a number of professional associations and consortia that publish corresponding specifications or recommendations on AI.[265] Many of the initiatives to tackle AI related challenges have been driven by industry, NGOs or consumer groups.

An important initiative to tackle AI challenges are the **Principles on Artificial Intelligence** developed by the **Organisation for the Economic Cooperation and Development (OECD).**[266] Published on 2019, the Principles aim to guide governments, organisations and individuals to design and run AI systems that are innovative and trustworthy and that respect human rights and democratic values. These principles are designed for AI systems generally, they are not specific to consumer products, and were developed by an OECD expert group on AI which includes representatives of governments, academia, professional organisations and businesses such as Facebook, Google, IBM and Microsoft.[267]

Similarly, at the European level, the **High-Level Expert Group on Artificial Intelligence (AI HLEG)** presented **the Ethics Guidelines for Trustworthy AI**,[268] which put forward a human-centric approach to AI and lists seven key requirements that AI systems should meet in order to be trustworthy. The AI HLEG is composed by 52 experts from science, civil society and industry and was appointed by the European Commission to provide advice and support on the implementation of the European AI strategy. This includes the development of recommendations for future policy development and ethical, legal and societal issues related to AI. The ethics guidelines for trustworthy AI cover topics such as fairness, security, transparency, future of work, democracy, privacy and protection of personal data.

The expert group has also presented a series of policy and investment recommendations; the AI HLEG made 33 recommendations to strengthen Europe's competitiveness, including guidelines for a strategic research agenda on AI and for the establishment of a network of AI centres of excellence. The recommendations aim to help the Commission and Member States to update their joint coordinated plan on AI.[269]

In addition, the Institute of Electrical and Electronics Engineers (IEEE) has been developing standards on AI for some years. The **IEEE's Ethically Aligned Design (EAD)**

[264] Hoichi Hamada et al (2019) Guidelines for Quality Assurance of Machine Learning-based Artificial Intelligence
[265] German Standardisation Roadmap
[266] OECD (2019) OECD AI Principles
[267] Simmons & Simmons (2019) OECD Principles on Artificial Intelligence
[268] AI HLEG (2019) Ethics guidelines for trustworthy AI
[269] DIN DKE (2020) German Standardization Roadmap on Artificial Intelligence

[270] document collects the inputs of several actors from academia, industry, policy and government and civil society on how to establish ethical and social implementations for intelligent and autonomous systems and technologies. The EAD identified a series of general principles (human rights, well-being, accountability, transparency and awareness of misuse[271]) as guidelines for the ethical design, development and implementation of these systems.[272]

Besides this initiative, the IEEE is also working to codevelop standards on other areas that include AI, such as personal data AI agent, transparency of autonomous systems, algorithmic bias considerations or wellbeing metric standards for ethical AI and autonomous systems.

The following common principles and values can be found in the OECD AI Principles, AI HLEG Ethics Guidelines for trustworthy AI and in the IEEE Ethically Aligned Design: transparency, security / safety, accountability, privacy, fairness and well-being.

At the UK national level, it is worth noting the work from the Alan Turing Institute and from the Centre for Data Ethics and Innovation (CDEI). The Alan Turing Institute, which is the national institute for data science and AI, undertakes research in collaboration with universities, business, and public and third sector organisations to tackle challenges related to science, society and the economy. In relation to AI, the Alan Turing Institute developed guidance designated to outline values, principles and the guidelines to assist department and delivery leads in ensuring that they develop and deploy AI ethically, safely, and responsibly.[273]

The CDEI is led by an independent board of expert members that advise the government on how the UK can maximise the benefits of the data-enabled technologies, including AI.[274] The CDEI highlights that AI can bring challenges that need to be managed, including the lack of the appropriate specialist knowledge to ensure that AI systems are trustworthy. The CDEI believes that an effective AI assurance ecosystem is required to address this information gap.[275] The CDEI is calling stakeholders in the public sector, industry, and academia to build insights into assurance tools and the assurance needs of those who rely on AI to draw out common insights and identify areas where clarity and consensus around AI assurance has the potential to bring about significant benefits.[276]

In addition, there are also a number of industry associations and organisations that are developing voluntary standards, codes of conduct, ethical principles and ethics frameworks for AI, for instance:

- The Software and Information Industry Association (SIIA) published the 'Ethical Principles for AI and Data Analytics' in which they highlighted that although AI can bring many benefits for consumers and other stakeholders, AI can also come with ethical challenges that should be addressed by policymakers and organizations

[270] IEEE (2019) Ethically Aligned Design- A vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems
[271] New technologies give rise to greater risk of misuse, and this is especially true for AI systems. AI increases the impact of risks such as hacking, the misuse of personal data, "gaming," or exploitation (e.g., of vulnerable users by unscrupulous parties). Responsible innovation requires designers to anticipate, reflect, and engage with users of AI
[272] European Commission (2018) State of the Art Report, Algorithmic decision-making, Algo:aware
[273] Leslie, D. (2018) The Alan Turing Institute – Understanding Artificial Intelligence ethics and safety. A guide for the responsible design and implementation of AI systems in the public sector
[274] Gowling WLG (2019) The Role of the Centre for Fata and Ethic and Innovation- What it Means for the UK
[275] Centre for Data Ethics and Innovation Blog (2021) The need for effective AI assurance
[276] Centre for Data Ethics and Innovation Blog (2021) Types of assurance in AI and the role of standards

themselves. The publication focuses on a framework of responsible data use that incorporates ethical principles for institutions to assess the data and models they use and to make modifications when they are needed. It also discusses the role of policymakers in addressing ethical concern.[277]

- The Information Technology Industry Council (ITI) advocates on behalf of their member companies and the larger technology ecosystem for policies that support the spirit of innovation while ensuring that AI remains inclusive and accessible. ITI has published a number of recommendations on national strategies on AI, including the EU's Strategy on AI.[278]

- Digital Catapult, a UK digital technology innovation centre that works with a range of organisations (e.g. startups, businesses, research and academia, public sector etc), also advocates for an ethical practice into the development of AI. It produced practical ethics guidance and recommendations for artificial intelligence technology development.[279]

- Underwriters Laboratories (UL) has developed voluntary standards for Safety for the Evaluation of Autonomous Products (ANSI/UL 4600). The standards address safety principles for fully autonomous systems that move, such as self-driving cars, along with applications in maintenance, agriculture, mining and other vehicles, including lightweight unmanned aerial vehicles (UAVs). Reliability of hardware and software necessary for machine learning, sensing of operating environment and other safety aspects of autonomy are also addressed. It is envisioned that future end-product standards will tailor UL 4600 to address specialized applications.[280] Although these standards do not relate to products within the scope of this report, it is interesting nonetheless that standards for specific products are starting to be developed.

[277] SIIA (2017) Ethical Principles for AI and Data Analytics
[278] https://www.itic.org/dotAsset/53157d6e-12cc-458d-bd97-f3d484237e14.pdf
[279] Digital Catapult - Digital Catapult unveils plan to increase adoption of ethics in artificial intelligence (digicatapult.org.uk)
[280] UL (2020) Presenting the Standard for Safety for the Evaluation of Autonomous Vehicles and Other Product

6 Framework of AI product safety policy considerations

This section presents a framework for considering the impact of AI consumer products on the product safety and liability policy framework. This framework aims to support the work of policymakers by highlighting the main considerations that should be taken into account when evaluating and developing product safety and liability policy for AI consumer products. This framework brings together the findings of all previous sections of this report.

First, it is key to understand the **nature of AI and consider the terminology used in relation to AI consumer products**. To summarise, there are simpler algorithms, for which the outputs are predictable, and more complex AI systems that produce outputs without being explicitly programmed. Currently, the former is more commonly found in consumer products, although the use of more complex AI systems will only grow in prevalence. However, the differences between product types and sectors also need to be considered. Furthermore, the picture of AI use in consumer products is clouded by conflation of the terms AI, smart, connected and IoT by all stakeholders.

Although the coverage of simpler applications of AI by the current product safety regulatory framework is clearer, there are a range of potential challenges brought by more complex applications. As such, before considering the possible product safety implications, it is useful to consider the following market-related questions:

- To what extent do consumer products use AI?
- What types of AI are being used, and for what purposes / functions?
- To what extent and how does this use differ by product type?

On this basis, it is important to **consider the unique ways in which AI consumer products can lead to product safety and liability challenges**. Building on the analysis of the characteristics of more complex AI consumer products, the following figure illustrates how these characteristics can lead to a range of potential challenges, which in turn can cause different types of harm.

Figure 6-1: Overview of AI-related product safety and liability challenges

The following table adds detail to the above figure, highlighting a range of considerations related to each AI characteristic.

Table 6-1: Key considerations, by AI characteristic

Characteristic (description)	Considerations
Mutability: Certain AI consumer products can learn and change over time.	• To what extent does the behaviour of an AI consumer product change over time? • To what extent does the existence and practical application of mutability differ by product type / sector?
Opacity: The workings of certain AI consumer products are opaque.	• To what extent are consumers aware of the use of AI in a consumer product (e.g. to make certain decisions)? • To what extent are the algorithms and models being used transparent and explainable?
Data needs: AI consumer products require a large amount of good quality data to function effectively.	• To what extent is good quality data available and accessible to economic operators? • To what extent and how is the data processed to minimise product risk?
Autonomy: AI consumer products can take decisions or actions without human intervention.	• To what extent are AI consumer products making decisions or taking actions autonomously? • To what extent are humans involved in governing the decisions and actions of an AI consumer product?

On this basis, it is important to consider the following two questions:

- **Challenges:** What impact does a characteristic, or a combination of characteristics, have on issues related to: robustness and predictability, transparency and explainability, security and resilience, fairness and discrimination, and privacy and data protection? Are there other possible challenges that do not fit in these areas?
- **Harms:** What potential harms can these challenges lead to? What is the likelihood of harms being caused? Are these potential harms material, immaterial or both? To what extent do the potential harms differ by product type / sector? To what extent are immaterial harms covered by other legislation?

Furthermore, the above challenges can occur as a result of decisions taken in the AI design and development process. As such, the role of the AI development process in ensuring the identified challenges are considered and tackled is also important.

As detailed through section 5, the **use of AI consumer products can also result in challenges to the regulatory framework for product safety and liability**. The following table summarises the challenges and related considerations.

Regulatory challenge (description)	Considerations
Product safety concepts and notions: The characteristics of AI and broader technological trends challenge the following regulatory concepts and notions: • Product • Safety and harm • Placing on the market • Producer	• To what extent are the characteristics of AI consumer products, as well as the possible challenges and harms, covered by the existing regulatory concepts and notions? • To what extent are these challenges solely related to AI as opposed to general technology and market trends?
Product liability concepts and notions: The characteristics of AI and broader technological trends challenge the following regulatory concepts and notions: • Product • Producer • Damage • Defect	• To what extent are the characteristics of AI consumer products, as well as the possible challenges and harms, covered by the existing regulatory concepts and notions? • To what extent are these challenges solely related to AI as opposed to general technology and market trends?
Regulatory mechanisms: The mechanisms supporting the legislation, including conformity assessment, market surveillance and standards, face challenges with regard to understanding and ensuring compliance of AI consumer products.	• To what extent do regulatory bodies (e.g. MSAs, conformity assessment bodies etc.) have the required skills and knowledge to assess, enable and enforce compliance of AI driven consumer products? • How can standards bodies best support compliance of AI consumer products?

In addition, certain **general market trends and other factors** are important to consider in relation to ensuring the effectiveness of the regulatory framework:

- **Complexities:** A range of factors related to the development and deployment of AI systems in consumer products increase the difficulty of ensuring the safety of products and, to that end, compliance with the regulatory framework. These factors include the increasing complexity of consumer products and the techniques used in AI applications, as well as the increasing complexity of the value chains necessary to develop AI driven consumer products. Furthermore, the ecosystems in which AI consumer products are being deployed are increasing in complexity, with interoperability and interactions between products ever more frequent. As such, regulators should consider the impact of the regulatory framework across these complexities, as well as in relation to isolated products or applications.

- **Coherence with other policy issues:** As highlighted throughout this study, the technical and policy challenges related to AI do not exist in a vacuum and are closely related to many other areas of technological and regulatory development. In particular, the following areas intersect with AI-related product safety challenges: challenges related to software updates and upgrades more broadly; cyber security

issues and related regulatory developments; environmental goals, including re-use and refurbishment; and (open) data policy, including data governance, as well as data protection.

Lastly, although there are clearly challenges related to the use of AI in consumer products, it is important to understand and consider that AI can also deliver significant **opportunities and benefits that AI**, both generally and for product safety specifically, as well as for all stakeholders. For instance, AI can support quality control for manufacturers, service diagnostics for users and compliance and enforcement for regulators. The above consideration of challenges needs to be understood in the context of these opportunities and benefits, as well as the need to ensure the regulatory approach enables innovation.

7 Conclusions

7.1 Terminology

Artificial intelligence (AI) is a broad term referring to computer systems that can sense their environment, think, possibly learn and take action in response to what they are sensing or their objectives.[281] Machine learning – the field of study that gives computers the ability to learn "without being explicitly programmed"[282] – is a key subset of AI. AI involves **mimicking intelligent human behaviours**, such as learning, prediction and adaptability based on immense amounts of data.[283] This can manifest, amongst other things, as **pattern recognition, image recognition, optimisation, and recommendation generation based** on data from a variety of media (videos, images, text, audio, etc.)[284]

However, **certain challenges persist regarding the use of the term artificial intelligence.** AI is often used as a buzzword in product marketing and is commonly conflated with related terms, such as 'smart' products, 'connected' products and consumer Internet of Things (IoT) products. As a result, the term AI is used to refer to a wide range of applications from quite simple algorithms to complex machine learning (ML) models.

7.2 Market for AI consumer products

In this context, it is **challenging to understand the true scale and dynamics of the market for AI powered consumer products.** Considering the size of the market, quantitative data exists on the scale of the market for consumer IoT devices, robotics and the total market for AI. Although these data suggest a continuously growing market, they do not specifically provide information on the scale of the AI consumer product market. Qualitative data collected through interviews and existing literature supports this general finding but indicates **notable differences between product groups.** While certain product groups, such as smart speakers, are found to be advanced in relation to the use of AI, other sectors, such as domestic appliances, currently report limited use of AI in existing products.

Although the use of AI in consumer products is found to be increasing, the research found several **barriers to adoption**. Primarily, these include cost, privacy and awareness.

In light of the challenges related to terminology, it is possible to identify some **key characteristics of AI applications** that are relevant in a product safety context. In particular, AI (and primarily ML) systems often need significant amounts of good quality data for training, testing and validation purposes; and they can be opaque on two levels, as: (i) it is often not clear to a consumer when an AI system is in use; and (ii) the workings of the technical approaches themselves can be opaque. In addition, AI systems often have the ability to learn and develop over time, instead of relying on explicit instructions, and they can display autonomy in their actions and decision-making.

---

[281] PwC. (2018). The macroeconomic impact of artificial intelligence.
[282] Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers, IBM Journal of Research and Development 44:1.2 (1959): 210–229.
[283] OII & Google. (2020). *Artificial Intelligence*. The A-Z of AI. [online]
[284] Ellen MacArthur Foundation. (n.d.). *Artificial Intelligence and the Circular Economy*. Ellen MacArthur Foundation. [online]

7.3 Opportunities and challenges for product safety

The **incorporation of AI systems into manufactured consumer products brings opportunities, as well as challenges and risks**. In terms of opportunities and benefits, there is a significant body of research highlighting the economic and social benefits of AI generally. When specifically considering product safety, the opportunities are also extensive, but can differ by product group. The direct opportunities for product safety include: more efficient and effective products; and predictive maintenance, which can directly improve product safety, as well as reduce maintenance costs and product downtime. In addition, indirect opportunities exist, including: improved data collection and analysis in the different phases of industrial assembly to increase product quality; improved cyber security protection; AI powered product design; and increasing potential for personalised products.

Considering the challenges and risks of AI to product safety, **the characteristics of AI as a technology highlighted above (including mutability, opacity, data needs, and autonomy) can translate into errors or challenges for AI systems that have the potential to cause harm.** These challenges can be categorised according to a range of themes, including robustness and predictability, transparency and explainability, security and resilience, fairness and discrimination, and privacy and data protection.

The **potential harms resulting from these challenges can be material or immaterial in nature**. Material harms, which are more likely to occur as a result of challenges in the first three themes (i.e. robustness and predictability, transparency and explainability, security and resilience), could include, for instance: an AI-driven robot malfunctioning as a result of automated decisions causing physical injury; or cyber security vulnerabilities in a product leading to threats to physical safety. Immaterial harms, which are more likely to occur as a result of fairness and discrimination or privacy and data protection challenges, could include, for instance, replacement of human contact for older people with autonomous products causing mental health issues; or discrimination in access to services for people with disabilities.

Beyond product safety risks specifically linked to AI, certain general trends can also bring product safety risks that can exacerbate or be exacerbated by AI consumer products. These include the tensions between built-in obsolescence and the circular economy, and the increasing reliance on e-commerce.

To date, however, **many of these risks are theoretical in nature and evidence of real-life examples of harm caused by AI consumer products is limited.** This most likely reflects a combination of factors, including: (i) the lack of maturity of many consumer product sectors in using AI; (ii) the existing consideration of the possible safety impacts of AI systems by the manufacturers and developers of these products; and (iii) the difficulty understanding the role and impact of AI systems when incidences do occur.

Beyond the potential impact of technical challenges on consumer harm, the **use of AI in consumer products can also challenge the regulatory framework for both product safety and liability**. For many existing AI consumer products, the current regulatory framework for product safety and liability and the mechanisms in place to monitor product safety are applicable and sufficient. However, the characteristics of more complex AI systems, in concert with general technological trends, pose challenges across all elements of the regulatory regime, including product safety and liability-related legislation, market surveillance regimes, standardisation, accreditation and conformity assessment. The key characteristics of AI systems, as highlighted above, include mutability, opacity, data needs and autonomy. The general market trends of relevance include: the blurring of the lines

between products and services; the increasing ability for consumer products to cause immaterial as well as material harm; the increasing complexity of supply chains for consumer products; and issues related to built-in obsolescence and maintenance throughout a product's lifecycle.

Considering the legislative framework for product safety and liability, more complex AI systems, as well as general technological and market changes, challenge many of the definitions detailed by these laws. More specifically, it is not clear to what extent these developments fall within the existing definitions of product, producer and placing on the market, as well as the related concepts of safety, harm, damages, and defects.

Furthermore, the characteristics of AI systems, the general trends highlighted, and the lack of clarity around the applicability of existing legal definitions and concepts, bring additional impacts. These include a lack of legal certainty for economic operators involved in the manufacture of AI driven consumer products, as well as a need to improve the skills and knowledge of regulatory bodies, such as MSAs and conformity assessment bodies, on AI systems.

Appendix A: Case studies on specific products

This appendix presents four case studies developed on specific product types: white goods, smart speakers, toys and robotics. The selection of case studies was conducted in collaboration with OPSS and focuses on some of the most common consumer products.

Case study: Use of AI in white goods

As seen earlier in this study, smart products are becoming more prominent in households across the UK, as consumers seek convenience, entertainment and more or improved features. Large domestic appliances, commonly known as white goods, are increasingly incorporating AI with the aim of facilitating household chores for consumers, saving them time and money. Examples of white goods include dishwashers, fridges, washing machines, microwaves and air conditioners. As mentioned, the term AI is often conflated with smart product. This case study aims to demonstrate the role of AI in a selection of products and the implications for safety.

Smart washing machines which have incorporated AI technology aim to improve the washing process. These washing machines use sensors to detect the volume, weight and fabric of each load. Using deep learning, these washing machines can triangulate this information with vast amounts of data to determine the optimal wash cycle, with the correct amount of detergent and customised motions, temperatures and times.[285] For instance, LG Turbowash uses AI DD™ to offer the most optimised wash based on 20,000 accumulated washes, or data points.[286] According to research, consumers are able to reduce detergent and power consumption by around 30% as a result of these capabilities.[287] Smart washing machines also offer consumers the option to start, pause or stop washing cycles with a smart phone, tablet or assistant. According to an interviewee from the industry, smart washing machines are at the higher end of the market but will likely become more popular as technology becomes more affordable.

The major concern related to these devices is security related, with consumers encouraged to update the apps from appropriate sources, choose appropriate app settings and avoid public Wi-Fi.[288] It can be argued, however, that smart washing machines are generally safer than conventional machines. An app allows the consumer to know if there are any issues with the machine, wherever they are; with a conventional machine, a consumer is unaware of its state after leaving the house and would not know if there had been a flood, for example. Data from the washing machine can be compared to a computer model; if trends or anomalies are present, predictions can be made to determine whether there might be a problem and preventative action can be taken. For example, if a prediction is made that the pump will break, the manufacturer can order the part, contact the owner and arrange a time to repair the machine.[289] Predictions can therefore prevent safety risks while cutting costs.

[285] https://www.forbes.com/sites/amandalauren/2020/01/20/how-lgs-artificial-intelligence-is-changing-how-consumers-use-appliances-improving-sustainability-and-doing-our-laundry-better/
[286] https://www.lg.com/uk/washing-machines/lg-F4V709WTS
[287] https://www.einfochips.com/blog/digitizing-homes-making-everyday-appliances-smarter-with-iot-and-ai/
[288] https://www.which.co.uk/reviews/washing-machines/article/smart-washing-machines-explained-atSzX5S9PxPE#smart-washing-machines-safety-and-security
[289] https://www.ibm.com/blogs/internet-of-things/washing-iot-solution/

This case study will now take a brief look at the smart kettle to show how a smaller domestic appliance works in comparison with larger products. A smart kettle allows the consumer to operate it remotely via an app or virtual assistant. Consumers can also control the temperature, check the water level and set times for the kettle to automatically boil.[290] Some smart kettles have incorporated volume sensing, whereby the kettle will start boiling upon hearing someone return home.[291] However, smart kettles are relatively new to market and there is limited research on them and their use of AI.

Case study: Use of AI in smart speakers

With many major tech and manufacturing companies producing voice-enabled smart speakers, these devices are renowned for their convenience, quality, accessibility and simple integration in the home. Their sophisticated natural language processing (NLP) abilities allow for clear conversations with their owners, while constantly improving their underlying functionality through machine learning.[292] However, the algorithms embedded in the cloud service these speakers connect to are constantly learning from the voices, commands, and conversations among their owners, eliciting privacy concerns.

According to techUK's annual State of the Connected Home report, smart speakers are more affordable than smart TVs, tablets and smartphones. The report also highlighted that, during the COVID-19 pandemic lockdowns, their voice-calling capabilities have allowed users to stay in touch with loved ones.[293] Indeed, one of the major benefits of smart speakers is that they allow owners to complete manual tasks such as cooking, building, repairing or gardening while using voice command to control instructional videos or radio programmes that are playing.[294] Their utility in these circumstances may have driven an increase in demand in recent years. Indeed, between 2019 and 2020, the proportion of UK citizens who own smart speakers increased from 22% to 29%, remaining second to only smart TVs in terms of ownership of smart products.[295] These figures are similar to Ofcom data, which in 2020 reported that 20% of individuals aged 16-64 had access to smart speakers.[296]

Despite the potential for smart speakers to become a hub for other IoT or AI-enabled devices in the home, according to the techUK report, smart speakers are currently more commonly used for entertainment (playing music, playing games) and information-seeking purposes (asking for instructions, checking the news, weather, or other information) rather than connecting to other smart home products.[297] When they are used for interoperability, they are most likely connected to a smart TV, smart lighting, smart plugs, smart thermostat, or an energy management service/app.[298] That final connection, albeit only at 3% of those surveyed, is an example of AI-enabled services communicating with one another.

---

[290] https://www.which.co.uk/news/2021/02/smart-kettle-explainer-what-you-need-to-know/
[291] https://littlehomeappliance.com/how-does-a-smart-kettle-work
[292] Cookson, M. (2020). *"Turn that on please" – Using Natural Interfaces in a half-engaged world*. Cambridge Consultants. [online]
[293] techUK. (2020). *The State of the Connected Home*. techUK. [online]
[294] Cookson, M. (2020). *"Turn that on please" – Using Natural Interfaces in a half-engaged world*. Cambridge Consultants. [online]
[295] techUK. (2020). *The State of the Connected Home*. techUK. [online]
[296] Ofcom. (2020). Adults' Media Use and Attitudes report, 20201/21. Ofcom.
[297] techUK. (2020). *The State of the Connected Home*. techUK. [online]
[298] techUK. (2020). *The State of the Connected Home*. techUK. [online]

The key barriers to adoption are cost and privacy. Smart speakers have garnered a negative reputation for listening, recording and storing its owners' audio data. This poses a privacy concern if these recordings are being used to improve the relevance of the adverts users see when browsing the Internet, or are simply heard by individuals who are not meant to be privy to private conversations. Although smart speakers are always listening, they are not always recording; recording commences once they hear their "wake word", which signals the owner is trying to make a command.[299] AI-enabled voice recognition means the speaker can differentiate between people within the home, and perhaps modulate its responses to a command based on how that user has behaved, or what they have asked for in the past. Even if the information is anonymous and encrypted, the recordings can still contain first names and other personal data, which is heard and noted by a transcriber.[300] Many consumers are simply uncomfortable with this level of intrusion, and not worth the convenience these speakers offer.

On an individual level, the level of distrust in smart speakers is significant. In a recent study on vulnerabilities in consumer IoT for DCMS, a survey conducted with consumers found that 57% of respondents believed it was "very likely" that a breach to consumer privacy via unauthorised access to devices such as smart speakers would occur, and 64% believed this would be "quite damaging". In terms of impacts in the wake of a cyber security incident, although no respondents reported an invasion of privacy via a smart speaker, they did state "time lost to resolving an issue" and "loss of trust in the brand/device/retailer" occurred.[301] This can also apply to any product safety issues that arise. Although the sample of respondents for this survey was not representative, the findings are illustrative of the concerns highlighted above.

In terms of regulatory challenges, one of the key issues is that smart speakers listen to and record everyone, including minors and other vulnerable groups. In the UK, it is legal to record conversations if done so for personal use. When these recordings are shared with third parties without the original participants' consent, it is considered an offence under the Regulation of Investigatory Powers Act 2000 (RIPA).[302] Indeed, user consent can become a persistent issue if there is little transparency on when the device is recording, how long it records, and where those recordings go.

Smart speakers have great potential, and are heralding a growing trend of voice command as a key feature of consumer AI products. However, privacy concerns must be considered, as well as users' comfort levels with corporations using their device interactions for commercial purposes.

Case study: Use of AI in toys

Connected toys comprise a diverse product category. Some pair a physical product with a smartphone app, such as drones, while others roam around the home and interact with users through voice command. AI plays a role in learning how to interact with its user, and react to external stimuli. Some are equipped with voice recognition technology

[299] Clauser, G. (2019). *Amazon's Alexa Never Stops Listening to You. Should You Worry?* New York Times: Wirecutter. [online]
[300] Clauser, G. (2019). *Amazon's Alexa Never Stops Listening to You. Should You Worry?* New York Times: Wirecutter. [online]
[301] CSES. (2020). *Framing the Nature and Scale of Cyber Security Vulnerabilities within the Current Consumer Internet of Things (IoT) Landscape*. CSES.
[302] Home Office. (2000). *Regulation of Investigatory Powers Act 2000 (RIPA).* Gov.uk. [online]

and NLP, while others have touch sensors.[303] As mentioned earlier in the report, one representative of the toy sector stated that much of the sector in the UK has withdrawn from developing these types of products, and it is unlikely the market will grow in the near future.

The benefits smart toys bring include introducing children to interacting with electronic devices and digital literacy from a young age, which can encourage them to be more adaptable to an increasingly connected world.[304] They can also bolster one's social development.[305]

However, the main concern surrounding these high-tech toys is privacy protection of minors. Toys are marketed primarily to children under 14, as stated earlier in this report, and therefore the data gathered and processed by the AI is that of minors. Since these devices respond to their owner's voices, they are constantly listening to conversations in the home; or, in conversations with their young owners, receive a litany of personal information that the child does not realise should be private. While listening is not the same as recording, the "wake" commands are intended to initiate conversations with the toy, at which point it begins recording. A prominent example that demonstrates these risks is the Cayla doll. The Cayla doll demonstrates how a toy with a paired smartphone app can send information via the app to the toy's manufacturers.[306] Furthermore, if a device is paired with an app, there may be personal details required to sign up for that service.[307] This risk has materialised as a barrier to product development and adoption, with manufacturers cautious about the risks and parents tending to divert their children away from screen-based and digital play.[308] Physical product safety concerns include whether the toy can still function if the manufacturer goes out of business, as many of these toys are produced by small or independent companies.[309]

Although there has been a general withdrawal among the toy sector from developing these products, major tech companies are still creating and selling AI-enabled products for children, even if the independent manufacturers are not. Therefore, it is still important to include toys in any product safety regulation for AI consumer devices.

Case study: Use of AI in robotics

Robotics are becoming increasingly prominent in consumer products, in line with developments in automation more generally. As seen already, the global market for personal and domestic service robots is growing, with sales expected to more than double by 2023. AI has driven many of the recent developments in robotics, enabling robots to sense and respond to their environment and perform an array of tasks benefiting the consumer. Robotics and AI are separate fields but share many characteristics; they can combine connectivity, autonomy and data dependency to carry out tasks and make decisions with little or no supervision. Many robots, however, are pre-programmed to carry out tasks which do not require AI. Algorithms, as discussed earlier, are key to performing complex tasks and improving performance.

[303] Tambini, O. (2018). *Are AI toys ethical?* Techradar. [online]
[304] Tambini, O. (2018). *Are AI toys ethical?* Techradar. [online]
[305] Kobie, N. (2015). *How smart are connected toys?* The Guardian. [online]
[306] Naylor, B. (2016). *This Doll May Be Recording What Children Say, Privacy Groups Charge*. NPR. [online]
[307] Maxwell, A. (2020). Robots, AI and drones: When did toys turn into rocket science?
[308] Cision. (2017). The UK Toys & Games Market 2017-2022.
[309] Maxwell, A. (2020). Robots, AI and drones: When did toys turn into rocket science?

Domestic robots have been in existence since the 1990s, assisting consumers with everyday chores, but have developed extensively in recent years. Robot vacuum cleaners, for example, are able to determine room sizes, adjust to carpet or hardwood, select the best routes and remember where objects are. A speech recognition AI engine can enable the robot to report the current status through verbal messages, while predefined messages and responses allow the robot to engage in a simple conversation with the consumer. These robots are continuously learning and adapting to their surroundings.[310] [311] Other examples of robots using AI are pool cleaning and window cleaning robots, entertainment robots which can interact, and domestic security robots.

Similarly, AI-enabled lawn mowers use the same principles. They map the field to ascertain where the grass needs to be cut and can target specific areas; they are also programmed to avoid objects. Moreover, they can analyse the temperature and rainfall to calculate when the grass should next be cut. Unlike robotic lawn mowers which do not rely on AI, the consumer is not required to mark the boundaries of the lawn.[312] [313] However, safety concerns remain. This case study will take a closer look at these devices.

A test conducted by German consumer safety group Stiftung Warentest demonstrated that not all robotic lawn mowers were able to precisely monitor their surroundings and avoid causing injury.[314] Some of the mowers were not able to stop in time when crawling children, represented by test dummies, were in the vicinity, leaving serious cuts. The test discovered that none of the devices stopped when a model of a child's hand was in front of them. Some of the mowers tested only rely on shock sensors, which are found a certain distance from the ground, and only detect and respond to obstacles on impact. The rotating blades can therefore damage a user's hands, arms or legs before they are identified.

To avoid the risk of injury, the Fraunhofer Institute for Microelectronic Circuits and Systems has developed highly advanced optic sensors which measure the surrounding area in 3D.[315] Incorporating LiDAR (Light Detection and Ranging) technology, they can measure how far away an object is as well as its speed based on the time it takes for an emitted light pulse to return after reflecting off the object. The 3D images generated distinguish between people and other objects, enabling the devices to keep clear of children or even shut down if a child approaches. Robotic lawn mowers serve as a good example of the safety risks and opportunities provided by robots incorporating AI. The device is constantly being developed with the aid of AI technology to enhance safety for consumers.

In recent years, Samsung Research has been developing AI-enabled companion robots to improve people's daily lives. 'Samsung Bot' is one of Samsung's next-generation AI projects which intends to aid users' physical activities and communicate with them through cognitive interactions. For example, it has developed 'GEMS (Gait Enhancing & Motivating System)', a walking assist wearable robot which helps those with weakened leg muscles enhance their physical performance. More features are being added, such as sport, fitness and entertainment functionalities, transforming living spaces for

[310] https://emerj.com/ai-sector-overviews/artificial-intelligence-home-robots-current-future-use-cases/
[311] https://ekosuunnittelu.info/wp-content/uploads/2019/06/Vacuum-cleaner-review_final-report-.pdf
[312] https://www.mobilegeeks.com/article/from-lawn-mowers-to-space-bosch-will-put-ai-in-everything/
[313] https://hypeandhyper.com/en/a-robotic-lawn-mower-powered-by-artificial-intelligence-toadi/
[314] https://www.fraunhofer.de/en/press/research-news/2018/august/increased-safety-for-children-around-lawn-mowers.html
[315] https://www.fraunhofer.de/en/press/research-news/2018/august/increased-safety-for-children-around-lawn-mowers.html

personal fitness needs. Samsung Research is aiming to achieve commercial-ready performance capabilities and verify and validate the safety and the accessibility for users.[316]

Another product developed by Samsung Research is Bot Care, which uses AI to recognise and understand behaviours to improve its performance as a robotic assistant. It knows the user's schedule and habits and can send reminders throughout the day. In addition, they are developing a robotic arm which utilises AI. Bot Handy is an Autonomous Mobile Manipulation Robot (AMMR) which acts as an extension of the consumer. It is able to recognise, grasp, pick up and place objects and can clean up untidy rooms and even sort out the dishes after a meal. The AI-enabled robotic arm comes with real-time control and indoor autonomous navigation technique.[317]

Similarly, Samsung has developed Bot Chef, an AI-powered collaborative robotic arm that can use everyday kitchen tools. Bot Chef can guide the user through recipe steps while cutting, mixing and seasoning as needed, and understands and communicates with the user. If the robot cannot find a particular item, it asks the user for help, and as soon as the user places the item within view of the robot it detects the object and works out how best to grasp it. Bot Chef has been designed with safety in mind. If it happens to get too close to the user while it is holding a sharp item, it either slows down or stops to keep the user's hands safe. In the case of impact, it is designed to merely bump into the user, not knock the user over.[318]

In healthcare, Medisana has teamed up with the company temi to develop The Home Care Robot which acts as a home health assistant. It provides comprehensive health monitoring, preventive healthcare and domestic independence. Utilising AI, it constantly learns and navigates autonomously throughout the home, and can be controlled by voice command or the touch display. An emergency call can be made to relatives by voice command or touch control in the event of an emergency, while up to four authorised relatives or nursing staff can remotely view the home and navigate the robot by remote control using the temi app. In future, automatic call forwarding to an external emergency service is planned, if personal contacts cannot be reached.[319]

316 https://research.samsung.com/robot
317 https://research.samsung.com/robot
318 https://cktinatinatina.medium.com/vp-of-research-at-samsung-talks-about-their-new-cooking-robot-samsung-bot-chef-interview-6125a77d27f5
319 https://www.medisana.com/en/The-smart-home-device-the-medisana-temi-Home-Care-Robot-celebrates-its-premiere-at-IFA-2019/

Appendix B: Bibliography

- AI Business. (2020). How smart speakers work.
- Algo:aware. (2018). State-of-the-Art Report | Algorithmic decision-making. Report developed for DG CNECT.
- Algorithm Watch, (2020), Automating Society Report 2020, United Kingdom.
- Amyx, S. (2015). Wearing Your Intelligence: How to Apply Artificial Intelligence in Wearables and IoT.
- Android.Developers. (2020). Distribution dashboard.
- Arnerić, S. P. et al. (2017). Biometric monitoring devices for assessing end points in clinical trials: developing an ecosystem. Nature Reviews Drug Discovery. 16(736).
- AT&T Foundry, Ericsson, & Rocketspace. (2018). The Future of Artificial Intelligence in Consumer Experience: According to the AT&T Foundry.
- Bayern, M. (2019). Manufacturers' digital transformation initiatives lag behind other industries. TechRepublic [online]
- BCG. (2018). Artificial Intelligence Is a Threat to Cybersecurity. It's Also a Solution.
- Becominghuman. (2019). How AI Can Improve Product Safety.
- BEIS. (2020). Consumer attitudes to product safety.
- BEIS. (2020). Smart Meter Statistics in Great Britain: Quarterly Report to end September 2020. BEIS.
- BEIS. (2020). The safety of domestic virtual reality systems.
- Bharadwaj, R. (2019). Artificial Intelligence in Home Robots – Current and Future Use-Cases.
- Bilodeau, S. (2019). Artificial intelligence in a "no choice but to get it smart" energy industry!
- Bobin, M., Amroun, H., Anastassova, M., Boukallel, M. & Ammi, M. (2017). In IEEE International Conference on Systems, Man, and Cybernetics.
- Brownlee, J. (2016). Overfitting and Underfitting with Machine Learning Algorithms, Article in Machine Learning Mastery. [online]
- Bryan, C. (2017). Amoral Machines, Or: How Roboticists Can Learn to Stop Worrying and Love the Law. Northwestern University Law Review, Vol. 111, No. 5, Fall 2017.
- BSI Standards. (2016). Robots and robotic devices. Guide to the ethical design and application of robots and robotic systems (BS 8611:2016).

- Cambridge Consultants, (2017), Review of the latest developments in the Internet of Things, study for Ofcom.
- CDEI. (2019). Snapshot Paper – Smart Speakers and Voice Assistants.
- CDEI. (2020). AI Barometer Report.
- CEN-CENELEC. (2020). CEN-CENELEC response to the EC White Paper on AI.
- CIO. (2018). 5 ways industrial AI is revolutionizing manufacturing.
- CloudTweaks. (2016). Digital Twin and the End of the Dreaded Product Recall.
- Consumer Affairs. (2019). The future of product recalls: AI and Amazon.
- Crawford, K. (2013). The Hidden Biases in Big Data, Harvard Business Review. [online]
- CSES and Tech4i2. (2020). Impact Assessment on Increased Protection of Internet-Connected Radio Equipment and Wearable Radio Equipment, Standalone Annex 8: Product-based case studies.
- CSES, (2020), Framing the Nature and Scale of Cyber Security Vulnerabilities within the Current consumer Internet of Things (IoT) Landscape, study for DCMS.
- CSES. (2020). Impact assessment on Increased Protection of Internet-Connected Radio Equipment and Wearable Radio Equipment, study for DG GROW, European Commission.
- CSES. (2020). Opportunities of Artificial Intelligence, study for the ITRE Committee, European Parliament.
- CSES. (2020). Study on the competitiveness of the EU engineering industries and the impact of digitalisation, study for EASME / DG GROW, European Commission.
- Daudu, A. (n.d.). These Headphones use Artificial Intelligence to protect your hearing.
- Deloitte. (2019). Future in the balance? How countries are pursuing an AI advantage.
- DeviceAtlas. (2019). Blog: Mobile OS versions by country.
- DigitalEurope. (2018). Recommendations on AI Policy Towards a sustainable & innovation friendly approach.
- EC-Council. (2019). Blog: The Role of AI in Cybersecurity.
- Edwards, L. (2016). Privacy, Security and Data Protection in Smart Cities, European Data Protection Law Review.
- Electronic Product Design and Test. (2019). Headphones gain AI-powered 'sense of hearing' via new Audio Analytic patent.

- Ellen MacArthur Foundation. (n.d.). Artificial Intelligence and the Circular Economy. Ellen MacArthur Foundation. [online]
- Emerj. (2020). Everyday Examples of Artificial Intelligence and Machine Learning.
- Energimyndigheten. (2017). Using webcrawler techniques for improved market surveillance – new possibilities for compliance and energy policy, presentation to the European Council for an Energy Efficient Economy (eceee).
- Esteva, A. et al. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature 542, 115–118.
- European Commission. (2014). 'Blue Guide' on the implementation of EU product rules.
- European Commission. (2020). Combined Evaluation Roadmap/Inception Impact Assessment: GPSD and AI.
- European Commission. (2020). Inception Impact Assessment: Proposal for a legal act of the European Parliament and the Council laying down requirements for Artificial Intelligence.
- Farhan, A.A. et al. (2016). Behavior vs. Introspection: Refining Prediction of Clinical Depression via Smartphone Sensing Data, IEEE Wireless Health (WH) (IEEE 2016).
- Forbes. (2019). The 10 Best Examples Of How AI Is Already Used In Our Everyday Life.
- Fowler, M. (1999). Refactoring: improving the design of existing code. Pearson Education India.
- Frank, P. (2016). The Black Box Society: the Secret Algorithms behind Money and Information. Harvard University Press.
- Gebhart, A. (2019). LG's new AI will tell you if you're using your fridge wrong.
- Giuffrida, I., Lederer, F., & Vermerys, N. (2018). A Legal Perspective on the Trials and Tribulations of AI: How Artificial Intelligence, the Internet of Things, Smart Contracts, and Other Technologies Will Affect the Law. 68(3), 747-780.
- Gröndahl, T. et al. (2018). All you Need is 'Love': Evading Hate-Speech Detection, Proceedings of the 11th ACM Workshop on Artificial Intelligence and Security (AISec) 2018.
- Hall, W., & Pesenti, J. (2017). Growing the artificial intelligence industry in the UK.
- IBM. (2016). How content analytics helps manufacturers improve product safety and save lives.
- IBM. (2019). Artificial Intelligence in Consumer Goods.
- IFC. (2020). Artificial Intelligence and the Future for Smart Homes.

- Intellectual Property Office. (2019). Artificial Intelligence, A worldwide overview of AI patents and patenting by the UK AI sector.
- International Federation of Robotics. (2020). Service Robots Record: Sales Worldwide Up 32%.
- International Federation of Robotics. (2020). World Robotics - Service Robots 2020.
- Irwin, B. (2018). Mass Customization of Personalized Digital Products.
- Kozyrkov, C. (2020) Training, validation and test phases in AI – explained in a way you'll never forget, Article in Towards Data Science. [online]
- Kumar, R., & Rasal, A. (2018). Smart Speaker Market by Intelligent Virtual Assistant, End User, Distribution Channel, and Price – Global Opportunity Analysis and Industry Forecast, 2018-2025.
- Laughlin, A. (2020). More than one billion Android devices at risk of malware threats. Which?
- Laurence, A. (2019). The Impact of Artificial Intelligence on Cyber Security.
- Layton, J. (n.d.). How Robotic Vacuums Work. Howstuffworks. [online]
- Lemley, J., Bazrafkan, S., and Corcoran, P. (2017). "Deep Learning for Consumer Devices and Services: Pushing the limits for machine learning, artificial intelligence, and computer vision.," in IEEE Consumer Electronics Magazine, vol. 6, no. 2, pp. 48-56, April 2017, doi: 10.1109/MCE.2016.2640698.
- Leslie, D. (2019). Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector. The Alan Turing Institute.
- Lorinc, J. (2019). Smart buildings: how AI is slashing heating and cooling bills.
- Masteron, A., Nahon, L. (2018) The UK's consumer product safety legal and regulatory regime. Pinset Mansons, Out-Law Guide
- Maurer, S., Punzano, F. (2020) BEUC and ANEC views for a Modern Regulatory Framework on Product Safety
- Maxwell, A. (2020). Robots, AI and drones: When did toys turn into rocket science?
- McCarthy, J. (2007). What is Artificial Intelligence?
- McKinsey Global Institute. (2019). Artificial intelligence in the United Kingdom: Prospects and challenges.
- McKinsey. (2017). Artificial Intelligence the Next Digital Frontier?
- McKinsey. (2020). The state of AI in 2020. McKinsey & Company. [online]
- Medeiros, J. (2018). How Voice Tech Is Slowly Including People With Speech Impediments.

- Mersinas, K., Sobb, T., Sample, C., Bakdash, J. Z., & Ormrod, D. (2019). Training Data and Rationality. In ECIAIR 2019 European Conference on the Impact of Artificial Intelligence and Robotics (p. 225). Academic Conferences and publishing ltd.
- National Consumer Federation. (n.d.) Benefit vs Risk in the new digital world.
- OECD. (2018). "Consumer product safety in the Internet of Things", OECD Digital Economy Papers, No. 267, OECD Publishing, Paris.
- OECD. (2019). Artificial Intelligence in Society, OECD Publishing, Paris.
- OECD. (2019). Challenges to consumer policy in the digital age.
- Ofcom, (2017), Connected Nations 2017: Data analysis, pp. 47-49.
- Ofcom. (2020). Online Nation: 2020 Report.
- OII & Google. (2020). Artificial Intelligence. The A-Z of AI. [online]
- OII & Google. (2020). Machine Learning. The A-Z of AI. [online]
- OPSS (2021) General Product Safety Regulations 2005: Guidance (GB)
- OPSS (2021) Supply of Machinery (Safety) Regulation 2008: Guidance (GB)
- OPSS. (2020). AI and product safety: summary of findings from the OPSS and Alan Turing Institute workshop.
- OPSS. (2021). UK Product Safety and Metrology: What's changed from 1 January 2021 in relation to Great Britain?
- Perc, M., Ozar, M., Hojnik, J. (2019) Social and juristic challenges of artificial intelligence
- PWC. (2017). The economic impact of artificial intelligence on the UK economy. PWC.
- PwC. (2018). The macroeconomic impact of artificial intelligence.
- RAPEX Alert Number: A12/0157/19 – Smart watch for children, 23/01/2019.
- Report from the Commission to the European Parliament, the Council and the European Economic and Social Committee. (2020). Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics, COM(2020) 64 final.
- Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers, IBM Journal of Research and Development 44:1.2 (1959): 210–229.
- Sculley, D. et al. (2015). Hidden Technical Debt in Machine Learning Systems, Google, Inc.
- Security Worldmarket. (2019). AI will transform the future of home security.

- Sharif, M. et al. (2016). Accessorize to a Crime: Real and Stealthy Attacks on State -of-the-Art Face Recognition, Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security.
- Stanford Institute for Human-Centered Artificial Intelligence (2019). The 2019 AI Index Report.
- Stanford University. (2016). Artificial Intelligence and Life in 2030.
- Tatman, R. (2017). Gender and dialect bias in YouTube's automatic captions. In Proceedings of the First ACL Workshop on Ethics in Natural Language Processing (pp. 53-59).BSI BS 8611:2016
- Taylor, L., Floridi, L. and van der Sloot, B. (2017) Group Privacy (eds) (Springer 2017).
- techUK. (2020). The State of the Connected Home.
- The Consumer Goods Forum & IBM. (2019). Artificial Intelligence in Consumer Goods.
- The General Product Safety Regulations 2005.
- Traverse for Citizens Advice. (2018). The future of the smart home: Current consumer attitudes towards Smart Home technology.
- Tschider, C. A. (2018). Regulating the internet of things: discrimination, privacy, and cybersecurity in the artificial intelligence age. Denv. L. Rev., 96, 87.
- Ustun, B. and Rudin, C. (2015) Supersparse Linear Integer Models for Optimized Medical Scoring Systems.
- Veale, M. and Binns, R. (2017). Fairer Machine Learning in the Real World: Mitigating Discrimination without Collecting Sensitive Data, 4 Big Data & Society 205395171774353.
- Weiser, M. (1994). The World is Not a Desktop, 1 Interactions 7.
- Which? (n.d.). Smart home security systems. Which? [online]
- Which? (n.d.). Wireless, smart and Bluetooth speakers. Which? [online]
- White Paper On Artificial Intelligence – A European approach to excellence and trust, COM(2020) 65 final.
- World Economic Forum. (2018). Here's how AI fits into the future of energy.
- Zeng, J., Ustun, B. and Rudin, C. (2017) Interpretable Classification Models f or Recidivism Prediction, 180 Journal of the Royal Statistical Society. Series A, 689.

Appendix C: Methodological approach and long list of AI topics

This section presents the methodological approach to the research and the long list of topics for in-depth research, submitted to OPSS as part of the interim report.

The methodological approach for this assignment comprised three phases and was conducted over a period of 21 weeks:

Phase 1: Preparatory phase

Project scoping activities were conducted in Phase 1, supporting the finalisation of the methodological approach to the study. These activities included a kick-off meeting with OPSS and a preliminary desk research exercise. The preparatory phase culminated in the delivery of the inception report, which was accompanied by data collection tools (see the interview topic guide in Appendix E). Feedback on the inception report was provided via an inception report review meeting.

Phase 2: Data collection & initial analysis

In the second phase, the study team conducted a desk research exercise and interview programme with the aim of collecting data relevant to all research questions and conducting initial analyses of those questions.

More specifically, the purpose of the **desk research exercise** was to identify and review literature and quantitative data sources to collect data and information relevant to all project objectives and research questions. As illustrated by the citations throughout and the bibliography (see Appendix B), the study team has reviewed a wide range of literature and data sources, including from UK, EU and global public authorities, industry and consumer associations, as well as academics and researchers. A 'Call for contributions' was also published on the CSES LinkedIn page.

Alongside the desk research exercise, the study team conducted an **interview programme**. The purpose of the interview programme was to collect data and perceptions relevant to all project objectives and research questions from representatives of all relevant stakeholder groups. A target of 30-40 interviews was set. The below table details the types of stakeholders consulted, the target number of interviews per group and the number of individual stakeholders from each group that were interviewed.

Interview programme

Stakeholder type (target)	Interviews completed
Academics, researcher institutes, think tanks & tech hubs (8-10 interviews)	6
Government / public bodies (4-5 interviews)	6
Law firms (1-2 interviews)	7
Manufacturers, AI developers & industry associations (12-15 interviews)	15

Stakeholder type (target)	Interviews completed
Product Safety practitioners & consumer associations (3-5 interviews)	6
Standards bodies, notified bodies & testing labs (3-5 interviews)	8
Grand Total	**48**

As highlighted in the above table, the study team contacted a total of 103 stakeholders and conducted interviews with 47 individual stakeholders spanning all six stakeholder groups.

In addition to the interviews conducted, the Smart Technology Product Safety Group, led by Electrical Safety First (ESF) and DLA Piper, submitted a written contribution in response to the call for contributions. Through correspondence with stakeholders, a range of literature was also signposted. This is reflected in the bibliography.

Analytical activities: In project Phase 2, the study team analysed the literature and data sources identified, alongside the interview feedback, to present initial assessments of the research questions under Objectives 1 and 2. This culminated in the submission of an interim report presenting the emerging findings on the: the regulatory framework for product safety; relevant definitions and terminology related to the use of AI in consumer products; the current and future market for AI-driven consumer products; the opportunities and challenges related to the use of AI in consumer products; and the related regulatory opportunities and challenges.

As requested by OPSS, the study team also developed a long list of specific topics, 5-8 of which would be selected for in-depth research in project phase 3. The long list of 19 topics is presented below. The following final selection of seven topics was agreed in collaboration with OPSS through the interim report review meeting:

- Mutability
- Robustness and predictability
- Transparency and explainability
- Impact on vulnerable consumer groups
- Immaterial harm
- Approaches to tackling AI risks in consumer products
- Liability

In addition to the above topics, it was agreed that five short case studies would be developed, covering the use of AI in four specific product groups (smart speakers, toys, robotics, white goods) and the challenges and opportunities brought by AI to market surveillance. The case studies on the four specific product groups can be found in Appendix A and the case study on market surveillance in section 5.1.

Phase 3: In-depth research and reporting

The purpose of the third and final project phase was to conduct further in-depth research on the seven selected topics, develop the five case studies, develop a framework of regulatory considerations and host a stakeholder workshop. For each topic and case study, the study team developed, in collaboration with OPSS, research questions to guide the in-depth research and analysis. On the basis of these research questions, the study team conducted a targeted literature review and analysis on each topic and case study. This analysis was presented in a draft final report.

Following the submission of the draft final report, a stakeholder workshop was held. Bringing together 29 representatives from all stakeholder groups, the study team presented the research findings before facilitating a discussion on the research issues, via two breakout groups. The research findings were subsequently presented to BEIS staff in a separate session.

The final project phase culminated in the submission of the final report, which incorporates all feedback received from OPSS, the stakeholder workshop and the presentation to BEIS staff.

Long-list of AI topics

The long list was developed based on the emerging findings presented in the interim report and contained 19 topics. In developing the long list of topics, a range of areas were considered, including: product groups, AI application types or functionalities, technical challenges or risks, regulatory challenges, and regulatory responses and debates.

Following the submission of the interim report, the long list of topics was discussed with OPSS. Seven topics and five case studies were selected for further research. As a result of overlaps and interlinkages between the topics, the final topics and case studies do not directly reflect the topics listed in the below table.

Summary of the long list of AI topics for further examination

AI topic: Description and rationale	Thematic area(s)
Monitoring and maintenance of AI systems post-market placement: Key questions will include: the extent to which the learning within ML systems is 'frozen' before being placed on the market or whether they continue learning; the role and mechanisms for upgrades and updates; and whether AI systems are located on the physical devices or if data is transferred and analysed in the cloud.	AI applications
Robotics and AI: Given the possibilities for AI applications in robots to result in physical harm to consumers, this topic would examine this type of product in more detail. Key questions would relate to: understanding what robotics refers to in the context of consumer products, the global regulatory approaches to regulating robotics and their application to AI (e.g. extensive considerations by the European Parliament and Japan).	AI applications
Predictive maintenance: This AI functionality is one of the more commonly used in consumer products. This topic would further explore how this is conducted in practice across different product groups and the benefits to safety.	AI applications

AI topic: Description and rationale	Thematic area(s)
Prominent product groups: Certain smart products, many of which use AI, are becoming widely adopted by consumers. Prominently, these include smart speakers and smart watches. This topic would further explore the role and implications of AI in these common products.	AI applications
Open-source AI: The use of open-source AI has the ability to support high-speed innovation and improve the resilience of code. This topic would explore the opportunities and challenges presented by open-source AI in the context of product safety. Furthermore, the prominent role played by large tech companies in the provision of open-source AI tools and methods could be examined.	AI applications
Barriers to AI development: Although certain product groups have made significant progress in using AI, many have not. This topic would explore the reasons why such sectors have not produced AI driven products. For instance, are there issues related to technical know-how or resources, is there no obvious consumer application. In this context, this topic could also examine the differences in the adoption of AI by large manufacturers and SMEs, including specific barriers related to knowledge and access to data / algorithms.	AI applications Technical challenges Regulatory challenges
Environmental applications: Given the policy focus on environmental issues, this topic would examine the environmental opportunities (e.g. energy efficiency) and challenges brought by AI consumer products. This could further examine smart meters and other energy sector products, as well as the tensions between built-in obsolescence and circular economy goals, re-use and recycling.	Technical challenges Regulatory challenges
Mutability: The inherent nature of ML models to learn and advance based on experience brings many technical and regulatory challenges. Further examination of these challenges could be useful, including questions such as how can ML products be scrutinsed prior to deployment or how can ML products be monitored and maintained over time.	Technical challenges Regulatory challenges
Data driven nature: AI applications in many instances also require a lot of data. Access to sufficient data, as well as unbiased and representative data, is a key challenge for implementing AI systems. This topic would further explore the impact this has on product safety.	Technical challenges
Robustness and predictability: Linked to the possible examinations of mutability and data, further research could be conducted on the challenges related to robustness and predictability. This assessment could focus on presenting a more holistic picture across the role of different AI characteristics in such challenges and a more detailed understanding of the possible outcomes for different product groups or application types.	Technical challenges

AI topic: Description and rationale	Thematic area(s)
Transparency and explainability: As above, transparency and explainability challenges can have impacts on product safety. This topic would further assess the possible mechanisms for tackling such challenges and debates around the standards to which decisions made by algorithmic systems should be held, as compared to human decision making. This topic would also further explore the impact of AI opacity on understanding how things go wrong and how liability can be attributed.	Technical challenges
Impact on vulnerable consumer groups: Many consumer products are used by vulnerable consumer groups or even explicitly designed for such groups (e.g. children, the elderly, people with disabilities). This topic could review in greater detail the possible product safety implications for different vulnerable groups of using AI in consumer products. This could consider both potential risks and harm, as well as access to the benefits of AI.	Technical challenges
Immaterial harm: It is being argued by some that product safety should also cover immaterial harm, including, for instance, mental health impacts of such products. Interesting questions here include the impacts of replacing human contact with products or increasing reliance on such smart consumer products. The coverage of privacy and data protection challenges related to AI consumer products could also be examined in greater detail.	Technical challenges Regulatory challenges
Concepts and definitions within the legislation: Linked to the above examination of immaterial harm, it could be useful to further examine the continuing relevance of certain concepts and definitions within the existing regulatory framework. For instance, beyond whether safety covers material and immaterial harm, this could focus on the concepts of product, producer, and placing on the market.	Regulatory challenges
Standardisation: Given the important role of standards in ensuring compliance with the current regulatory framework, this topic could take a further look at ongoing standardisation activities. This could examine work at the UK (BSI), EU (CEN-CENELEC) and international (ISO, IEEE) levels related to AI, looking at the types of standards being developed and their relevance to the use of AI in consumer products.	Regulatory responses
Liability: To date, the research has focused on product safety and the related legislation. However, there are clearly impacts related to attributing liability. As such, this topic would further examine the UK framework for liability and the possible impact of AI in the context of consumer products.	Regulatory responses
Market surveillance: Although not necessarily AI specific, market surveillance authorities are facing many challenges related to new technologies. For instance, the increasing use of e-commerce. Furthermore, market surveillance authorities, as well as conformity assessment bodies, face challenges related to the skills needed to	Regulatory responses

AI topic: Description and rationale	Thematic area(s)
assess AI consumer products. On the other hand, AI can bring significant benefits to market monitoring and recall effectiveness.	
Industry / non-legislative approaches: To date, many of the initiatives to tackle AI related challenges have been driven by industry, NGOs or consumer groups. This topic could examine the nature and success of such approaches, including codes of conduct, ethical guidelines and practical ethics and AI safety tools. Under this topic, we could examine industry approaches to self-regulation, as well as the assessment list of EU High-Level Expert Group on AI and other work by international organisations.	Regulatory responses
International regulatory developments: As highlighted throughout the report, the EU has taken a range of steps towards regulating AI generally as well as in the context of product safety. This topic would examine their ongoing actions in more depth, as well as those of other leading countries, such as China and the US.	Regulatory responses

Appendix D: List of stakeholders consulted

This appendix presents a list of stakeholders consulted through the project, via interviews and/or attendance at the stakeholder workshop.

Type	Organisation
Academics, researcher institutes, think tanks & tech hubs	Cambridge University
	Institute for Ethical AI, Oxford Brookes University
	King's College London
	The Alan Turing Institute
	Trustworthy Autonomous Systems Hub, Governance and Regulation Research Node
	Trustworthy Autonomous Systems Hub (King's College London)
Government / public bodies	Centre for Data Ethics and Innovation (CDEI)
	Danish Safety Technology Authority
	Office for AI
	UK Research & Innovation (UKRI) / Innovate UK
Law firms	Allen & Overy
	Bird & Bird
	CMS
	Cooley LLP
	Freshfields
	Law Commission
Manufacturers, AI developers & industry associations	Amazon
	Association of Manufacturers of Domestic Appliances (AMDEA)
	BEAMA (UK Trade Association for Manufacturers and Providers of Energy Infrastructure Technologies and Systems)
	Beko
	Developers Alliance

	Google
	Huawei
	Sony
	techUK
Product Safety practitioners & consumer associations	Citizens Advice
	Electrical Safety First
	National Consumer Federation
	Trading Standards (CTSI)
Standards bodies, notified bodies & testing labs	British Standards Institution (BSI)
	European Committee for Standardization (CEN) and European Committee for Electrotechnical Standardization (CENELEC)
	TUV SUD
	UL International (UK)
	United Kingdom Accreditation Service (UKAS)

Appendix E: Interview Topic Guide

Introduction

- Please provide an overview of your organisation and role.

For manufacturers / industry associations: What type of products does your organisation manufacture? To what extent do these products use AI? What types of AI do your products use (e.g. machine learning, deep learning, natural language processing etc.)?

- Do you have any introductory remarks or questions related to this research project?

Objective 1: Current and likely future applications of AI in the home

1. What does the current and forecasted future market of AI-driven consumer products look like? Please comment in relation to the following aspects:

   - Market size, structure, supply chain dynamics
   - Current consumer trends (i.e. what AI products are consumers buying)
   - Future consumer trends and whether anything has changed in the past 5-10 years
   - Types of AI being used in consumer products (e.g. natural language processing; machine learning; pattern recognition; which types of algorithms etc.)
   - Sectors / product types that are most impacted by AI use

Are you aware of any quantitative data sources relevant to these elements?

2. What are the key similarities and differences between AI and non-AI products and how can AI be defined when applied to products?

3. How and to what extent do you think AI technologies and their application to products will change in the next 5-10 years? What will be the main drivers of these changes?

4. What product safety opportunities does the use of AI in consumer products introduce or enhance?

   - How might these opportunities evolve in the next 5-10 years?
   - Examples of the realisation of such opportunities.

5. What product safety challenges and risks does the use of AI in consumer product introduce or enhance?

   - Which features of AI are contributing to these product safety challenges and risks?
   - How might these challenges and risks evolve in the next 5-10 years?
   - Examples of the realisation of such challenges and risks.

Objectives 2 & 3: Current product safety framework and factors affecting regulation

6. What characteristics of AI should be considered when regulating consumer products? (e.g. ability to keep learning; reliance on software updates).

7. How might the characteristics of AI change how harm and safety is addressed by regulation? (e.g. physical vs. non-physical harm).

8. What are the regulatory challenges stemming from the incorporation of AI systems in consumer products?

9. To what extent is the existing regulatory framework for product safety effective in ensuring product safety in consumer products that use AI?

10. Are there any regulatory gaps in the UK's current AI safety framework?

11. How can AI systems be used to support the regulation of product safety?

12. What possible policy options exist to respond to these product safety challenges and fill any regulatory gaps, while facilitating and fostering product innovation? What are the possible impacts of these policy options?

13. Are you aware of any successful strategy or best practices to tackle the regulatory challenges posed by AI to product safety? In the UK, or the rest of the world?

Contact us if you have any enquiries about this publication, including requests for alternative formats, at: OPSS.enquiries@beis.gov.uk

Office for Product Safety and Standards

Department for Business, Energy and Industrial Strategy
4th Floor, Cannon House, 18 The Priory Queensway, Birmingham B4 6BS
https://www.gov.uk/government/organisations/office-for-product-safety-and-standards