ARTIFICIAL INTELLIGENCE ENTITIES AND CRIMINAL LIABILITY: A NIGERIAN JURISPRUDENTIAL DIAGNOSIS[1]

Abstract
Contemporary technological inventions are beginning to support or replace human activities with the emergence of artificial intelligence entities ranging from autonomous cars to machines translation software, robots and medical diagnosis software. These inventions tend to venture into some human mental activities such as interpretation, evaluation, and decision-making, which have never been delegated to non-human mind before. Sometimes, criminal activities are associated with these Artificial Intelligence Entities. Generally, criminal liability arises basically out of the presence of two factors: mens rea and actus reus. Actus reus is the physical outcome of the act. In the case of artificially intelligence being, the main challenge arises in dictating the mens rea which is the mental factor. This paper evaluates sundry views and examines whether human laws can be imposed on the artificial intelligence entities like they are imposed on legal personalities like corporations whereby punishments can be awarded by making necessary alterations. This study recommends, among others, the application of the Turing Test by the court to detect whether the artificial Intelligence entity is capable enough to formulate mens rea. Qualitative methods are used in the analysis of data collected from text books, journals and internet resources.

Keywords: *Artificial Intelligence, Criminal liability, Mens rea, Jurisprudence, Nigeria*

1. Introduction

Human intelligence is a natural endowment to man. With this intelligence, man seeks to create entities that would further the manifestation of his intelligence. Yet, more and more simple human activities are being replaced by robots and computers. Artificial intelligence is becoming an integral part of human society with all its implications, negative and positive. In 1981, a 37-year-old Japanese employee of a motorcycle factory was killed by an artificial-intelligence robot working near him. The robot erroneously identified the employee as a threat to its mission, and calculated that the most efficient way to eliminate this threat was by pushing him into an adjacent operating machine. Using its very powerful hydraulic arm, the robot smashed the surprised worker into the operating machine, killing him instantly, and then resumed its duties with no one to interfere with its mission. Unfortunately, this is not science fiction, and the legal question is: If advanced autonomous machine commits a crime of its own accord, how would it be treated in law? How would the guilty mind of a non-human be demonstrated and how can it be done within existing criminal law principles? Can a crime be said to have been committed? Indeed, is the machine liable given the principles of criminal liability? This study will examine some models set out set out by scholars for determining criminal liability when an Artificial Intelligence entity is involved in a crime. These views will be juxtaposed with the present state of Nigerian laws with effect to examining their propriety or otherwise in the light of the growing use and prospects of Artificial

[1]***Ikenga K.E. ORAEGBUNAM**, **PhD (Law), PhD (Phil.), PhD (Rel. & Soc.), MEd, BTh, BL,** Senior Lecturer and Ag Head, Department of International Law and Jurisprudence, Faculty of Law, Nnamdi Azikiwe University, P.M.B. 5025, Awka, Anambra State, Nigeria. Email: ikengaken@gmail.com; ik.oraegbunam@unizik.edu.ng. Phone Number: +2348034711211; and
***Uguru Eme UGURU, LLB, BL, LLM Candidate,** Faculty of Law, Nnamdi Azikiwe University, P.M.B. 5025, Awka, Anambra State, Nigeria.

Intelligence entities and systems in Nigeria. The study will be capped with the prospects of the use of Turing Test in determining whether or not machines and robots are endowed with the capacity to develop guilty or innocent minds.

2. Artificial Intelligence: Attempts at Definition

Since Artificial Intelligence is a new discipline, which covers an extensive field, it has no single definition.[2] Development and research of Systems of Artificial Intelligence began with the emergence of the first personal computers.[3]John McCarthy[4] introduced the term 'Artificial Intelligence' in 1955, defining it as 'the science and engineering of making Intelligence machines, especially Intelligence computer programs'. However, Artificial Intelligence is a broad concept with myriad applications from the 'Intelligence' assistants in our cell phones (i.e. Siri® of Apple®) to 'Intelligence' home appliances to any future technologies that may cause a paradigm shift in our understanding of life.[5] Even though the use of Systems of Artificial Intelligence is becoming more and more extensive in the everyday lives of many people, it is still rather difficult to come up with an accurate definition of the concept.[6]Systems of Artificial Intelligence covers wide fields of processes from basic tasks, such as legal reasoning, to specific ones such as playing chess, checking mathematical theories, etc.

In their studies, various authors[7]define Artificial Intelligence as artificially developed intelligence; for example, a software system that is able to imitate human ways of thinking with the help of a computer or other device (house management systems integrated in household devices, robots, self-driving cars, etc.). Artificial Intelligence is the capability of a machine to imitate Intelligence behavior.[8]Artificial Intelligence is the simulation of human behavior and cognitive processes on a computer and hence is the study of the nature of the whole space of Intelligence minds.[9]Cognitive processes use existing knowledge to generate new knowledge. Since human cognition or cognitive processes encompass processes such as acquiring knowledge and understanding through thoughts, experiences and the senses, as well as processes like attention, memory and working memory, judgment, and evaluation, reasoning and computation, problem solving, decision making, comprehension and production of language, it therefore means that creating a software or machine with Artificial Intelligence is an attempt to imitate all these cognitive processes in man and vest them in a machine as much as possible. Artificial Intelligence research began in the 1940s and early

[2] Gonenc Gurkaynak, Ilay Yilmaz, Gunes Haksever; 'Stifling artificial intelligence: Human perils' *computer law & security review* 32 (2 0 1 6) p. 749–758.
[3] ibid
[4]American computer scientist and cognitive scientist, professor at Stanford University, Massachusetts Institute of Technology,Dartmouth College and Princeton University. Died on October 24, 2011.
[5]Gonenc Gurkaynak, Ilay Yilmaz, Gunes Haksever (n. 2).
[6]Paulius C˘erka, Jurgita Grigiene˙, Gintare˙ Sirbikyte; 'Is it possible to grant legal personality to artificial intelligence software systems?' *computer law & security review* 33 (2017) p. 685–699.
[7]Stuart Russell, Peter Norving, *Artificial Intelligence: A Modern Approach*(3rd edn, NJ Prentice Hall 2009) 4-5
[8]Gabriel Hallevy; 'The Criminal Liability of Artificial Intelligence Entities—From Science Fiction to Legal Social Control' *Akron Intellectual Property Journal* 4:171, p. 172
[9] ibid

1950s. Since then, Artificial Intelligence entities have become an integral part of modern human life, functioning much more sophisticatedly than other daily tools.[10]

There are five attributes that one would expect an Intelligence entity to have.[11]The first is communication[12].One can communicate with an Intelligence entity. The easier it is to communicate with an entity, the more Intelligence the entity seems. The second is internal knowledge.[13] An Intelligence entity is expected to have some knowledge about itself. The third is external knowledge.[14] An Intelligence entity is expected to know about the outside world, to learn about it, and utilize that information.[15]The fourth is goal-driven behavior.[16]An Intelligence entity is expected to take action in order to achieve its goals.[17] The fifth is creativity.[18]An Intelligence entity is expected to have some degree of creativity. In this context, creativity means the ability to take alternate action when the initial action fails.[19]Some twenty-first century types of Artificial Intelligence entities possess even more attributes that enable them to act in far more sophisticated ways.[20]

3. Manifestation of Artificial Intelligence

An Artificial Intelligence entity has a wide variety of applications, including in robots. Artificial Intelligence robots and Artificial Intelligence software are used in a wide range of applications in industry, military services, medical services, science, and even in games.[21] In order to provide a better understanding of what Artificial Intelligence means when it comes to reality, it has been found worthwhile to classify the 'intelligence level' of Artificial intelligence under three main groups: Artificial Narrow Intelligence (ANI), Artificial General Intelligence (AGI) and Artificial Super Intelligence.[22] From this classification we are able to give examples of forms in which Artificial Intelligence entities manifests in our environment to have a better grasp of what they are. ANIs are Artificial Intelligence Entities specialized in a specific area, such as IBM's DeepBlue®, the supercomputer that beat Gary Kasparov, the reigning World Chess Champion in May 1997.[23] As with most ANIs, Deep Blue's abilities were limited.[24]Playing chess was the only thing Deep Blue® could do.[25] ANIs surround us today and they work in a similar manner.[26]The Intelligence thermometers of 'Nest®', Apple's 'Siri®', videogames, search engines, social networks, web

[10] ibid
[11]Gabriel Hallevy (n. 8), p. 175-176
[12] ibid
[13] ibid
[14] ibid
[15] ibid
[16] ibid
[17] ibid
[18] ibid
[19] ibid
[20] ibid
[21]Gabriel Hallevy (n. 8), p. 177
[22]Gonenc Gurkaynak, Ilay Yilmaz, Gunes Haksever (n. 2).
[23] ibid
[24] ibid
[25] ibid
[26] ibid

cookies, online advertising services, data miners and data scrapers, autopilots, traffic control software, automated phone answering services and so on; neither of which can initiate thought processes in order to provide queries falling outside the scope of their predetermined operations.[27] This is the level of Artificial Intelligence humanity has achieved so far.[28] However, ANI's talents are constantly getting better and more impressive. Speech recognition and processing allows computers to convert sounds to text with greater accuracy. Google® is using Artificial Intelligence to caption millions of videos on YouTube®.[29]

AGIs represent 'Human-Level AIs', computers as smart as humans in every aspect and capable of performing all intellectual tasks humans can. AGIs are expected to be capable of solving various complex problems in various different domains with the ability of autonomous control with their own thoughts, worries, feelings, strengths, weaknesses and predispositions.[30] ASIs represent AIs 'much smarter than the best human brains in practically every field, including scientific creativity, general wisdom and social skills.'[31] ASIs are the AIs that many fear, will optimize Earth, aiming to fulfill their goals, by eventually removing human kind from the face of it.[32]The majority of AI scientists foresee that, after the development of an AGI, it will evolve itself into an ASI very quickly, as a result of an exponential growth loop. This phenomenon is also known as an 'intelligence explosion' or ''singularity'.[33]

From a legal point of view, the physical or non-physical appearance is less important than the different degrees of autonomy or intelligence these agents possess, which range from human supervision (level 1), and deterministic autonomy (level 2), to machine-learning (level 3) and multi-agent systems (level 4). Some agents operate independently, but are supervised by humans who may intervene or give clearance for certain actions. This level 1 autonomy is also labeled 'human *on* the loop'. Autonomous cars or autonomous weapon systems, for example, are currently still being supervised by a human who can redirect the car or has to authorize an attack. Here, the operator is a 'human *on* the loop' as opposed to 'human *in* the loop' who retains control over device function. Agents operating without human supervision or 'human *out* of the loop' further differ, even though all 'loop' operations rely on algorithms (ie step-by-step problem-solving procedures). Deterministic algorithms (level 2) solve problems the designer foresaw and provide designer-anticipated responses. For example, an algorithm designed to translate traffic laws into computer code would prevent a self-driving car from crossing a solid center line on the road. More advanced algorithms are capable of 'learning' (level 3) which means that algorithms improve behavior through experience (i.e. by training, not by following a fixed program). Algorithm could for example learn to recognize road users (cars, cyclists, pedestrians) or traffic signs after being fed with images depicting those users or signs. This process involves analyzing data, detecting patterns (for example shapes, colours or positions of objects) and building or refining models, mostly based on statistical calculations in order to fulfil a given task. It is a particularly useful method when providing all the

[27] ibid
[28] ibid
[29] ibid
[30] ibid
[31] ibid
[32] ibid
[33] ibid

steps necessary for fulfilling a given task is complicated (as in language translation), when these steps are unknown (as in detecting the risk of a certain disease), and when large data sets exist or may easily be created (as in discovering spending habits of internet users). Consequently, machine-learning allows algorithms to detect problems and solutions their programmers did not foresee. Machine-learning can also be important for a meaningful translation of law into computer code. A self-driving car could learn from model scenarios or from watching the other road users, for example, that it can ignore the rule not to cross a solid center line in exceptional cases allowing the car when it detects debris in its lane, for instance, to cautiously circumnavigate the obstacle instead of bringing traffic to a standstill. As this example already indicates, machine-learning also comes with different autonomy. 'Supervised learning' indicates that a human monitors the learning process more closely, typically by classifying the training data (e.g. specifying the legal and illegal behavior of road users). 'self-learning' however, allows for an open process of learning and reaction (for example if a self-driving car simply imitates the behavior of other road users.[34]

Yet, dynamics get more complex if artificial intelligent interact with one another to coordinate their behavior (level 4). Self-driving cars again provide illustrative examples here. In order to navigate safely, speedily, and ecologically, autonomous cars will not only communicate with digital road infrastructure ('V21'), but also with other road users ('V2v'), receiving and providing information on circumstances such as road conditions, traffic flow or parking space. Upon receiving the relevant information, cars will respond by adapting their route (driving around a bottleneck) or their speed (optimizing the traffic flow). The adequate reaction may be prescribed by a central actor, such as a traffic agency in charge of the road infrastructure. But it may also be found in de-centralized manner, like amongst a group of vehicles in proximity to each other. Thus, two cars may coordinate which one of them will go first if the right of way is unclear in a particular situation; several cars may coordinate their speed and direction of travel respectively to avoid traffic congestion.3 Artificial intelligent entities taking part in decentralized forms of decision making may not be more autonomous than those described above. However, from the perspective of human control and supervision, their actions are even less predictable. Machine-learning can prompt self-driving cars to misclassify objects or to violate the law; interaction with other artificial agents might provoke unintended results. The work will concern itself with the responsibility for an artificial agent's decisions and actions that were not intended or foreseen by its creator, programmer or user.

It is, however, apt to note that performing tasks that involve complex calculations requiring substantial effort, time and dedication for humans are very simple for Artificial Intelligence Systems.[35] However, tasks that seem so simple for us, such as voice and image recognition, movement, anticipation and perception are extremely hard for Artificial Intelligence Systems, or the coders of Artificial Intelligence Systems.[36] Hence, it has been opined that 'Artificial Intelligence Systems have by now succeeded in doing essentially everything that requires 'thinking' but has failed to do most of what people and animals do 'without thinking', that, somehow, is much harder!'[37]

[34]https://ssrn.com/abstract/artificial agents and general principles of law accessed on 22nd of April, 2018
[35] Gonenc Gurkaynak, Ilay Yilmaz, Gunes Haksever (n. 2).
[36] ibid
[37] ibid

4. Notion of Criminal Liability

The basic question of criminal law is the question of criminal liability, that is, whether the specific entity (human or corporation) bears criminal liability for a specific offence committed at a specific point in time and space.[38] There is a lot of points to be noted in the laws guiding criminal liability in Nigeria. A perusal of chapter 5 of the Criminal Code, especially sections 22 – 36 thereof, reveals these. However for the purpose of this study, we will limit them to a few points which are of immediate relevance. The first is the provision of section 36 (12) of the Constitution of the Federal Republic of Nigeria which states that 'a person shall not be convicted of an offence unless that offence is defined and the penalty therefore is prescribed in a written law'. The next is Section 28 and 30 of the Criminal Code which exempts from criminal liability persons which the law deems to lack the mental capability (*doli incapaxi*) of committing an offence whether by reason of tender age or insanity. Section 30 exempts from criminal responsibility for any act or omission all persons under the age of 7 years, and also persons under the age of 12 years for any act or omission, unless it is proved that at the time of doing the act or making the omission he had capacity to know that he ought not to do the act or make the omission. Section 28 is to the effect that one is not criminally responsible for an act or omission if at the time of doing the act or making the omission the person is in such a state of mental disease or natural mental infirmity as to deprive him of capacity to understand what he is doing or of capacity to control his actions, or of capacity to know that he ought not do the act or make the omission.

Again of relevance to this topic is the provision of Section 7 of the Criminal Code Act which provides for parties to offences and who principal offenders:

> When an offence is committed, each of the following persons is deemed to have taken part in committing the offence and to be guilty of the offence, and may be charged with actually committing it-
> (a) every person who actually does the act or makes the omission which constitutes the offence;
> (b) every person who does or omits to do any act for the purpose of enabling or aiding another person to commit the offence;
> (c) every person who aids another person in committing the offence;
> (d) any person who counsels or procures any other person to commit the offence.
> In the fourth case, he may be charged either with himself committing the offence or with counseling or procuring its commission.

Lastly, there is need to mention the universal elements or ingredients in fixing criminal liability on a person. This was developed in the common law and came to be part of Nigeria's jurisprudence although no criminal statute in Nigeria spells it out. It is expressed in the Latin maxim, actus *non facit reum nisi men sit rea*, which means that – an act does not render a person legally guilty unless his mind is also blameworthy. This maxim has evolved into the two elements or ingredients for fixing criminal liability which are *Actus Reus* (the guilty act) and *Mens Rea* (the guilty mind). In other words, the first is the external or factual element—i.e., criminal conduct (*actus reus*)—while the other is the internal or mental element—i.e. knowledge or general intent vis-à-vis the conduct

[38]Gabriel Hallevy (n. 8), p. 177

element (*mens rea*).[39] Generally, once one element is missing, no criminal liability can be imposed. The *actus reus* requirement is expressed mainly by acts or omissions. Sometimes, other external elements are required in addition to conduct, such as the specific results of that conduct and the specific circumstances underlying the conduct. The *mens rea* requirement has various levels of mental elements. The highest level is expressed by knowledge, while sometimes it is accompanied by a requirement of intent or specific intention.[40] Lower levels are expressed by negligence (a reasonable person should have known) or by strict liability offences.[41]No other criteria or capabilities are required in order to impose criminal liability, not from humans, nor from any other kind of entity, including corporations and Artificial Intelligence entities.[42]An entity might possess further capabilities; however, in order to impose criminal liability, the existence of *actus reus* and *mens rea* in the specific offence is quite enough.[43]

At this juncture, certain questions and responses are apt. The first is: When is a crime said to be committed. In Nigeria as with most parts of the world, a crime is said to be committed when a person recognized as such by the law and who is not statutorily excluded from being criminally culpable, does an act or makes an omission defined by the statutes to be an offence, and such a person did such an act or made such an omission with the required criminal knowledge or intent. When this happens, a crime can be said to have been committed. A follow up question then is; whether a machine or an Artificial Intelligence Entity can be said to have committed a crime so as to be criminally liable? To attempt to answer this, copious reference will be made to the three liability models set by Gabriel Hallevy[44] used to determine who bears liability for acts or omissions of an Artificial Intelligence entity so as to know who the law sees as actually having committed the crime. This study also examines the relevant provisions of Nigerian law on criminal liability to know if as they stand, a machine can be imputed with a crime. Paramount is the issue of how Artificial Intelligence entities can fulfill the two requirements of criminal liability.

Gabriel Hallevy proposes the imposition of criminal liability on Artificial Intelligence entities using three possible models of liability: the Perpetration-via-Another liability model; the Natural-Probable-Consequence liability model; and the Direct liability model.[45] The Perpetration-via-Another liability model does not consider the Artificial Intelligence entity as possessing any human attributes.[46] The Artificial Intelligence entity is considered an innocent agent.[47]Accordingly, due to that legal viewpoint, a machine is a machine, and is never human. Although one cannot totally ignore an Artificial Intelligence entity's capabilities, as mentioned above, according to this model, these capabilities are insufficient to deem the Artificial Intelligence entity a perpetrator of an offence.[48] These capabilities resemble the parallel capabilities of a mentally limited person, such as a child, a

[39] ibid
[40] ibid
[41] ibid
[42] ibid
[43] ibid
[44]Gabriel Hallevy (n.8) p. 179
[45]Gabriel Hallevy (n. 8) p. 179-183
[46] ibid
[47] ibid
[48] ibid

person who is mentally incompetent, or one who lacks a criminal state of mind.[49] Legally, when an offence is committed by an innocent agent (a child, a person who is mentally incompetent, or one who lacks a criminal state of mind to commit an offence), the intermediary is regarded as a mere instrument, albeit a sophisticated instrument, while the party orchestrating the offence (the perpetrator-via-another) is the real perpetrator as a principal in the first degree and is held accountable for the conduct of the innocent agent.[50]The perpetrator's liability is determined on the basis of the 'instrument's' conduct and his own (the perpetrator's) mental state. The derivative question relative to artificial Intelligence entities is: Who is the perpetrator-via-another?[51]There are two candidates: the first is the programmer of the Artificial Intelligence software and the second is the user, or the end-user.[52] Is it then the programmer who designs the Artificial Intelligence entity to commit an offence or an end-user who although did not design the Artificial Intelligence entity uses it for his own benefit and commits an offence in the process. In both cases, the actual offence was committed by the Artificial Intelligence entity. Of course the programmer or the user did not perform any action conforming to the definition of a specific offence; therefore, none meets the *actus reus* requirement of the specific offence.[53]The perpetration-via-another liability model considers the action committed by the Artificial Intelligence entity as if it had been the programmer's or the user's action.[54]No mental attribute required for the imposition of criminal liability is attributed to the Artificial Intelligence entity.[55]When programmers or users use an AI entity instrumentally, the commission of an offence by the AI entity is attributed to them. The internal element required in the specific offence already exists in their minds.[56]When an end-user makes instrumental usage of an innocent agent to commit a crime, the end-user is deemed the perpetrator.[57] This liability model does not attribute any mental capability, or any human mental capability, to the AI entity.[58] This is since according to it, there is no legal difference between an Artificial Intelligence entity and a screwdriver used instrumentally by a burglar to open a window pane or an animal used for similar purposes. The screwdriver or animal is not criminally liable. The screwdriver's 'action' is, in fact, the burglar's. Be that as it may, the perpetration-via-another liability model is not suitable when an AI entity decides to commit an offence based on its own accumulated experience or knowledge.[59] This model is not suitable when the software of the AI entity was not designed to commit the specific offence, but was committed by the AI entity nonetheless.[60] This model is also not suitable when the specific AI entity functions not as an innocent agent, but as a semi-innocent agent.[61] The legal result of applying this model is that the programmer

[49] ibid
[50] ibid
[51] ibid
[52] ibid
[53] ibid
[54] ibid
[55] ibid
[56] ibid
[57] ibid
[58] ibid
[59] ibid
[60] ibid
[61] ibid

and the user are criminally liable for the specific offence committed, while the AI entity has no criminal liability whatsoever.[62]

The Natural-Probable-Consequence liability model assumes deep involvement of the programmers or users in the Artificial Intelligence entity's daily activities, but without any intention of committing any offence via the Artificial Intelligence entity.[63] For example, during the execution of its daily tasks, an Artificial Intelligence entity commits an offence. The programmers or users had no knowledge of the offence until it had already been committed; they did not plan to commit any offence, and they did not participate in any part of the commission of that specific offence.[64]For such cases, the second model might create a suitable legal response. This model is based upon the ability of the programmers or users to foresee the potential commission of offences.[65]According to this model, a person might be held accountable for an offence, if that offence is a natural and probable consequence of that person's conduct.[66]Originally, the natural-probable-consequence liability model was used to impose criminal liability upon accomplices, when one committed an offence, which had not been planned by all of them and which was not part of a conspiracy.[67]The natural-probable-consequence liability model requires the programmer or user to be in a negligent mental state, not more.[68] Programmers or users are not required to know about any forthcoming commission of an offence as a result of their activity, but are required to know that such an offence is a natural, probable consequence of their actions.[69]The programmers or users of an AI entity, who should have known about the probability of the forthcoming commission of the specific offence, are criminally liable for the specific offence, even though they did not actually know about it.[70]Reasonable programmers or users should have foreseen the offence, and prevented it from being committed by the AI entity.[71]

However, the legal results of applying the natural-probable-consequence liability model to the programmer or user differ in two different types of factual cases.[72] The first type of case is when the programmers or users were negligent while programming or using the AI entity but had no criminal intent to commit any offence[73]. The second type of case is when the programmers or users programmed or used the AI entity knowingly and willfully in order to commit one offence via the AI entity, but the AI entity deviated from the plan and committed some other offence, in addition to or instead of the planned offence.[74] The first type of case is one of pure negligence.[75]Consequently,

[62] ibid
[63] ibid
[64] ibid
[65] ibid
[66] ibid
[67] ibid
[68] ibid
[69] ibid
[70] ibid
[71] ibid
[72] ibid
[73] ibid
[74] ibid
[75] ibid

if there is a specific offence of negligent homicide in that legal system, this is the most severe offence for which the programmer might be held accountable, not manslaughter or murder, which requires knowledge or intent.[76]In the second type case, the programmer shall be held criminally liable for both the offence it originally programmed it to commit and also for the offence it committed in addition to or instead of the original programming. The question still remains: What is the criminal liability of the AI entity itself when the natural-probable-consequence liability model is applied?[77] In fact, there are two possible outcomes. If the AI entity acted as an innocent agent, without knowing anything about the criminal prohibition, it is not held criminally accountable for the offence it committed.[78] Under such circumstances, the actions of the AI entity were not different from the actions of the AI entity under the first model (the perpetration-via-another liability model).[79] However, if the AI entity did not act merely as an innocent agent, then, in addition to the criminal liability of the programmer or user pursuant to the natural-probable-consequence liability model, the AI entity itself shall be held criminally liable for the specific offence directly.[80]

The direct liability model of AI entities is the third model, as described hereunder. The third model does not assume any dependence of the AI entity on a specific programmer or user. The third model focuses on the AI entity itself.[81] As noted earlier, criminal liability for a specific offence is mainly comprised of the external element (*actus reus*) and the internal element (*mens rea*) of that offence.[82] Any person attributed with both elements of the specific offence is held criminally accountable for that specific offence.[83] No other criteria are required in order to impose criminal liability. In order to impose criminal liability on any kind of entity, the existence of these elements in the specific entity must be proven.[84]When it has been proven that a person committed the offence in question with knowledge or intent, that person is held criminally liable for that offence[85]. The relevant questions regarding the criminal liability of AI entities are: How can these entities fulfill the requirements of criminal liability?[86] Do AI entities differ from humans in this context?[87] If an AI entity is capable of fulfilling the requirements of both the external element and the internal element, and, in fact, it actually fulfills them, there is nothing to prevent criminal liability from being imposed on that AI entity.[88]If a person fulfills the requirements of both the external element and the internal element of a specific offence, then the person is held criminally liable.[89] Why should an AI entity that fulfills all elements of an offence be exempt from criminal liability?[90]This is in brief the contention of the direct liability model.

[76] ibid
[77] ibid
[78] ibid
[79] ibid
[80] ibid
[81] ibid
[82] ibid
[83] ibid
[84] ibid
[85] ibid
[86] ibid
[87] ibid
[88] ibid
[89] ibid
[90] ibid

To be criminally liable, artificial intelligence entities need to be treated as legal persons just as corporations are legal persons under the law. Of course the initial reasoning behind according legal personhood to corporations was to promote commercial activity and also remove corporate liability from individual shoulders. In the same vein, artificial intelligence should be accorded basic constitutional freedoms in line with those accorded to corporations. The primary objective behind this is that as artificial intelligence develops and begins to think, civil and criminal liability arising from their actions will not be solely attributable to their programmer or owner. Take for instance the issue of the autopilot based on the artificial intelligence technology. What if a developer of a warfare aircraft makes an autopilot program which itself eliminates any obstacles on its mission and in one of the mission the pilot of the aircraft aborts the mission due to bad weather but the autopilot recognizes pilot as an obstacle and ejects the pilot out of the cabin which kills the pilot. Now the developer did not have any intention to kill the pilot but the current laws consider them liable. The correct option would be to impose criminal liability on the autopilot and correct the algorithms of its programming. This not only saves the developers of artificial intelligence and the owners from criminal liability for acts they never intended but also prohibits demoralization of developers from bringing more innovations.[91] However, the crime perpetrator can easily take shelter behind artificial intelligence beings and use legal personality of the artificial intelligence entity as a statutory privilege to commit crimes. In the case of a corporation if any person uses legal personality of the corporation for his fraudulent of dishonest purposes he is not allowed to take shelter behind the legal personality of the corporation and the court lifts the corporate veil of the corporation and takes action against the perpetrator as there is not corporate personality. The corporate veil is lifted only if a person relies on the corporate personality of the corporation to shield his wrong doings. In the same vein, the scenario of artificial intelligence can be treated. If a perpetrator of any fraud or crime is found taking shelter behind the legal personality of the robot he should be treated by a court as if there was no legal personality. Many precedents of which are being slowly established like the case of 'computer raped by telephone' which was widely reported in which a programmer used a telephone link to invade the privacy of the computer. During the course of the investigation the questions arose as to whether a search warrant can be issued to the computer to fetch evidences.[92]

It has been said, however, that the combined and coordinated application of the three models above reveals a new legal situation in the specific context of AI entities and criminal law.[93] As a result, when AI entities and humans are involved, directly or indirectly, in the perpetration of a specific offence, it will be far more difficult to evade criminal liability. The social benefit derived from such a legal policy is of substantial value. All entities, human, legal, or AI become subject to criminal law.[94] If the clearest purpose of the imposition of criminal liability is the application of legal social

[91] A. Rafael, Calvo, Sidney K. D'Mello, Jonathan Gratch & Arvid Kappas, *The Oxford Handbook of Affective Computing* (1st edn, Oxford 2015) 176

[92] *BSN (UK) Ltd v Janardan Mohandas Rajan Pillai* [1996] 86 Com Cases 371 (Bom)s. This was the first time when the world saw any computer being treated as a person and a search warrant was issued to the computer.

[93] ibid

[94] ibid

control in the specific society, then the coordinated application of all three models is necessary in the very context of AI entities.[95]

Attempts have been made to face the challenge of detecting *mens rea* in the acts of the robots by the application of Turing test. In 1950, Alan Turing introduced the concept of Turing Test to test the ability of a machine to formulate intents for its actions or to exhibit intelligence.[96] Turing test is a game in which the machine imitates being a human with a human opponent. After a series of questions, the questioner who is completely unaware of which competitor is human and which one is a computer guesses which of them is human. Turing test basically tests the ability of a machine to exhibit human nature. If the machine is successful in convincing the questioner that it is human, it passes the test and is believed to have capabilities to act as a human. The applicability of the Turing test on every particular artificial intelligence entity can be cumbersome for the courts. Therefore, the government can lay down norms for manufacturers or developers of these entities to subject them to Turing test or any other test as the government deems fit before the public offering of these entities. Alan Turing's Test was passed for the first time in 2014 as This- go-round, a Russian-made programme, which disguised itself as a 13-year-old boy named Eugene Goostman from Odessa, Ukaine, bamboozled 33 percent of human questioners. Eugene was one of five supercomputers who entered the2014 Turing Test. Kelvin Warwick, a visiting professor at the University of Reading, who organized the event at the Royal Society in London noted that 'in the field of Artificial intelligence there is no more iconic and controversial milestone than the Turing Test, when a computer makes a sufficient number of interrogators to believe that it is not a machine but a human'.

5. Artificial Intelligence Entities and Liability for Crimes in Nigeria

How can Nigerian law on criminal liability be juxtaposed with offences committed by or with the use of Artificial Intelligence entities? Do Nigerian criminal laws cover this development? The word 'person(s)' in Nigeria laws defining offences and setting out punishments is relevant. Section 36 (12) of the Constitution of the Federal Republic of Nigeria 1999 (as amended) clearly uses the word 'person' in enshrining one of the principles of legality that must be considered before anyone is tried or punished for any alleged or purported criminal offence. This shows that it is a sine qua non that before criminal liability is imputed, it must be certain on whom the imputation is being made. The question is, is the person a proper subject recognized by law? A failure to get this right will lead to futility when it comes to sentencing and practically executing punishment. Artificial Intelligence entities are not anywhere in Nigerian laws recognized as having legal personality to be subjects of the law.

Criminal law embodies the most powerful legal social control in modern civilization.[97] People's fear of Artificial Intelligence entities, in most cases, is based on the fact that Artificial Intelligence entities are not considered to be subject to the law, specifically to criminal law.[98] In the past, people were similarly fearful of corporations and their power to commit a spectrum of crimes, but because

[95] ibid
[96] Alan Turing, *Computing Machinery and Intelligence* [1950] LIX 236
[97]Gabriel Hallevy (n. 8) p. 173-174
[98] ibid

corporations have become legal entities subject to criminal and corporate law, that kind of fear has been significantly reduced.[99] This same result can be achieved with Artificial Intelligence entities. It will not only obviate the general fear for them and for their criminal tendencies; it will also make the laws certain and the combined effect of these is that Artificial Intelligence entities will revolutionize Nigeria- just like how corporation have imparted national economies and Gross Development Product greatly over the past century. A lot is being done through corporations because it has been clothed with flesh and blood by the law. A lot more can be done for humanity and growth in Nigeria if same is done to Artificial Intelligence entities.

There are reasons for giving legal personality to Artificial Intelligence entities. It has been proposed that it is impossible to ignore the idea that as we come to interact with robots and Systems of Artificial Intelligence as though they are persons, we will increasingly grant them legal rights.[100] If Society accepts agents in the regular commercial transactions as legal entities, which are recognized subjects of law, so why could Systems of Artificial Intelligence not be granted the same status?[101]These systems possess the elements characteristic of subjects that have been granted legal personality: SAI are Intelligence, able to make autonomous decisions, learn from their own experience – memorize, plan, demonstrate complexity, formality and ability to manipulate structures, which let them interact with other subjects of law.[102] Another reason to recognize Systems of AI subjects of law is their immediate interaction with other subjects of law, optimum protection of whose rights and interests requires clearly defining the legal status of Systems of AI.[103]However, presently under the current legal regulation at national and international level, Systems of AI are seen only as objects of law.[104]

Another issue that can be drawn from the provision of our law in section 36(12) of the Constitution is that if legal personality is not granted Artificial Intelligence Entities, whether crimes that can be or will be committed can be covered by the already defined offences in Nigerian penal books. In other words is there no need to extend the meaning of specific offences like murder, stealing etc. to cover novel cases where AIs or their end users or programmers commit them. The recommendation therefore is that this can be achieved by making new legislation or amendment of old ones. This was the case in Nigeria when scamming went cyber and the offences of fraud and corruption in the Criminal Code could not properly cover the new situations that were arising daily from cyber fraud. Cybercrimes (Prohibition, Prevention, Etc) Act 2015 constitutes a response to this development. In fact, with AIs there is great possibility that new crimes not formerly known to Nigerian law can be committed by AIs or with the aid of AIs, hence the need for prompt legislative reaction.

The next issue relates to the applicability of sections 28 and 30 of the Criminal Code. Do AIs have mental capacity so as to be held liable for an offence? Since it is generally believed that machines are excluded from criminal liability for lack of mental capacity to know the nature of their act as one that would result in an offence (since it is the programming of the maker or command of the

[99] ibid
[100]Paulius C˘ erka, Jurgita Grigiene˙, Gintare˙ Sirbikyte (n. 6) p. 685–699.
[101] ibid
[102] ibid
[103] ibid
[104] ibid

user and not theirs) or to form or have general or specific intent (since it can be argued that machines do not know good or bad except the commands given it). Thus, the mental element to make them criminally liable is missing. Again, if per adventure Nigerian laws confer on AIs legal personality as persons and proper subjects of the law, will there be room to classify some AI entities as minors by age and experience (having learnt that many AIs are programmed to grow over time in their human simulated cognitive processes by continuously gathering knowledge from their experiences gleaned from doing the acts they had been programmed to do, somewhat akin to infant humans growing in mind to adulthood and maturation). Again would there be room to classify some AIs as mentally infirm or diseased and to exclude them from criminal liability under Nigerian law by virtue of this mental infirmity which can be a malfunction that occurred in its programming causing it to do an act which it was never built or intended to do? In other words, leaving the laws the way they are without more, can sections 28 and 30 of the Criminal Code be construed to apply to AIs or be extended to them? This is what gives rise to the dire need for legislation on AI in Nigeria to avoid wrong construction of laws not intended for such purposes.

Another area that can be examined from the provisions of Nigerian law on criminal liability in relation to Artificial Intelligence entities is on the provisions relating to parties to an offence. In section 7 of the Criminal Code, certain parties to an offence are stated as principal offenders and as liable to the same degree as the actual doer or omitter of the act, regardless of their level of involvement. All the subsections can be construed to cover cases of crimes committed by AIs or through the agency of AIs. They can also be linked to the three liability models for imputing criminal liability on AI entities discussed above. For instances, subsection (a) can be used to hold either the AI itself culpable i.e Direct Liability Model- or even to hold the programmer or end user culpable, for getting the AIs physical acts and taking them as that of the programmer or end user.[105] Subsection (b) and (c) can be used to hold the end user or programmer criminally liable. Although they did not actually know when the offence was committed nor did they plan it, they should have reasonably foreseen it and prevented it. Hence, their act or omission in not reasonably foreseeing and hence aiding the AI entity itself to do an act or omit to do an act resulting in an offence will be imputed on them. This is akin to the Natural- probable- Consequence liability model as it is a model for accomplice liability cases. Subsection (d) can be used to hold criminally liable a person be it the programmer or the end user who uses the AI as an agent or instrument to perpetrate a crime. In the words of the Criminal Code, it can be said that such end users or programmers procured another for the commission of the offence or counseled another for the commission of the offence. The Code makes them liable for not just counseling or procuring but for also doing the actual act or making alleged omission. This is the way one can construe what is available in Nigerian law to meet the Perpetrator-via-agent liability model.

6. Conclusion

Ultimately, it can be confidently said that the base for the growth of Artificial Intelligence and Artificial Intelligence technology has already taken shape and is forging ahead as we grapple daily to understand and acquaint ourselves with the new forms it takes. Yet, Nigerian laws cannot be said to be certain applying them the way they are in relation to this sphere of human endeavor. It leaves room for a lot of construction and stretching which may lead to the law reaching its breaking point and being incapable of covering the desired grounds. The solution therefore is for Nigerian laws to recognize Artificial Intelligence Entities as legal personalities and subjects of law. This will then open up the frontiers for the required legislative intervention to cover the scope of Artificial Intelligence systems in relation to criminal liability.

[105] See above for the exact provisions of the identified subsections.