eLaw
Working Paper Series

No 202ť/00 - ELAW– 202ť

MU,uz,O~u y uzmx,Xm¤

Mz, [£q~£uq¤,{r,MU,M| |xuom¡u{z ,uz,
_¢n ¡mz¡u£q,mzp,\~{oqp¢~mx,O~u y uzmx,Xm¤
O¢ ¡q~ ß,N:T:Y:,

Universiteit Leiden
eLaw

Discover the world at Leiden University

Chapter 11
AI in Criminal Law: An Overview of AI Applications in Substantive and Procedural Criminal Law

Bart Custers

Contents

Abstract Both criminals and law enforcement are increasingly making use of the opportunities that AI may offer, opening a whole new chapter in the cat-and-mouse game of committing versus addressing crime. This chapter maps the major developments of AI use in both substantive criminal law and procedural criminal law. In substantive criminal law, A/B optimisation, deepfake technologies, and algorithmic profiling are examined, particularly the way in which these technologies contribute to existing and new types of crime. Also the role of AI in assessing the effectiveness of sanctions and other justice-related programs and practices is examined, particularly risk taxation instruments and evidence-based sanctioning. In procedural criminal law, AI can be used as a law enforcement technology, for instance, for predictive policing or as a cyber agent technology. Also the role of AI in evidence (data analytics after search and seizure, Bayesian statistics, developing scenarios) is examined. Finally, focus areas for further legal research are proposed.

B. Custers (✉)
eLaw–Center for Law and Digital Technologies at Leiden University, Leiden, The Netherlands
e-mail: b.h.m.custers@law.leidenuniv.nl

B. Custers and E. Fosch-Villaronga (eds.), *Law and Artificial Intelligence*,
Information Technology and Law Series 35,
https://doi.org/10.1007/978-94-6265-523-2_11

Keywords cyber agent technology · deepfake technologies · evidence-based sanctioning · law enforcement technology · predictive policing · risk taxation instruments

11.1 Introduction

Artificial Intelligence (AI) is the new hype. In many countries, large amounts of funding are available for further research on AI.[1] It may be expected that AI will bring significant changes in several sectors of society, including transport (e.g., self-driving cars), healthcare (e.g., automated drug discovery), education (e.g., adaptive virtual tutors catering to personalized individual needs), and language (e.g., real-time translations of conversations). Also in the legal domain AI is expected to bring change. On the one hand, developments in AI may call for new, different or further regulation and, on the other hand, AI may offer more and more applications for legal research and legal practice.[2] This chapter aims to provide an overview of AI developments in the area of criminal law, both in substantive criminal law and procedural criminal law. When discussing substantive criminal law, this chapter focuses on the use of AI by criminals and the use of AI when imposing sanctions or other justice-related programs. When discussing procedural criminal law, the focus of this chapter is on the use of AI in criminal investigation and prosecution and the role of AI in criminal evidence. In both parts it is investigated which new (types of) legal questions these developments raise.

All examples used and described in this chapter are real, existing examples, not future or hypothetical examples. Furthermore, this chapter does not include a section defining what AI is and which technologies can be considered AI. No clear definition of AI exists in literature and at points there is even a lack of convergence on what AI exactly is.[3] To steer clear of this debate on what counts as AI and what not, this chapter only discusses AI technologies that are self-learning and autonomous. Most of the AI discussed in this chapter is technology based on machine learning.[4]

This chapter is structured as follows. Section 11.2 discusses developments in substantive criminal law and Sect. 11.3 discusses developments in procedural criminal law. Section 11.4 provides conclusions and identifies focus areas for further legal research.

---

[1] Rosemain and Rose 2018. In the US: Harper 2021. In the Netherlands: https://nlaic.com and https://www.universiteitleiden.nl/en/sails. Many countries worldwide are thinking about developing policies for AI, see Jobin et al. 2019, pp. 389–399.

[2] Custers 2018, pp. 355–377.

[3] Calo 2017.

[4] Calders and Custers 2013.

11.2 AI and Substantive Criminal Law

11.2.1 Crimes

Developments in AI technologies such as data mining and machine learning enable several new opportunities for criminals to commit new types of crimes and new ways of committing well-known crimes. This section describes several examples of new types of crime enabled by AI and types of crime that have are rapidly becoming more prevalent due to the use of AI. Most of the technologies used for the applications discussed in this section, such as A/B optimisation, are based on data mining and machine learning and focus on the discovery of patterns. Deepfake technologies are not necessarily focused on pattern discovery, but also are a form of deep learning, usually based on artificial neural networks.

11.2.1.1 A/B Optimisation

Many websites, such as online stores, websites for booking hotel rooms, and news websites, use so-called *A/B testing*[5] (also referred to as A/B optimisation). A/B testing means that some visitors to the website are offered screen A (or version A) and other visitors get screen B (or version B). Version A and B only have one difference, sometimes very subtle. For instance, the difference can be black versus dark blue text colours, or the background colour is pale yellow instead of pale blue, or the headers in the text are underlined in one version, but not in the other version. Both versions are then monitored in terms of how long visitors stay on the website, click on advertisements, or order something. If version A turns out to yield better results than version B, the latter version is rejected and the former is continued with. By repeating this many times and offering different versions to large numbers of visitors, an optimized result can be achieved. In fact, all internet users are used as guinea pigs to find out what works best.[6]

Obviously, this A/B testing is not a manual procedure—it is automated and usually self-learning, which makes it a form of AI. Usually it is algorithms (based on technologies like data mining and machine learning) that discover particular patterns. Self-learning software can also create on its own these variations in the lay-out of a website or the text in a message. Via algorithmic decision-making, the information is then offered to the users in a specific way. It is important to stress that A/B testing does not require any personal data. It can also be applied to anonymous visitors of a website and it is not a form of personalisation. It is about general preferences, not about personal preferences.

Companies can use A/B testing to retain people longer on their websites, supporting the attention economics, and even to increase the number of product

[5] Kohavi and Thomke 2017, pp. 74–82.

[6] Gallo 2017.

sales. Also criminals use this approach. When criminals start using phishing (i.e., trying to obtain bank account details of their victims), ransomware (i.e., trying to lock computers or files of their victims and order a ransom), or WhatsApp fraud (i.e., trying to convince their victims to transfer money to a friend in need), the challenge for the criminals is always the same: convincing a victim to click on a link or an attachment that will install malware or, even more directly, to transfer money.[7] In other words, criminals are always looking for the most convincing screens. The use of A/B testing and many guinea pigs can help achieve this. The spam used in all these types of cybercrime is not simply a free trial (like the term phishing suggest), it also offers criminals to watch and see what works (i.e., when victims take the bait) and optimise their methods (like the term spear phishing expresses).[8]

As a result of these developments, the fake screens we see look increasingly real. Distinguishing what is real and what is fake becomes more and more difficult, for instance, for messages from a bank or employer. In WhatsApp fraud, for instance, often profile pictures of friends or family members are used to increase the trustworthiness of messages. It is not surprising that unsuspecting victims fall into these traps in increasingly large numbers. Europol reports a rise in these types of cybercrime year after year for several years now.[9]

11.2.1.2 *Deepfake Technology*

Related to this, there is another AI technology that deserves attention. Deepfake technology offers the possibility of merging images and videos. It is also possible to generate completely new footage, for instance, of non-existing people through AI.[10] This technology is cheap and little technological knowledge is required. Deepfake technology can make someone look better or worse, or even completely different, as is shown in Fig. 11.1. In some cases, deepfake technology can merge pictures or videos of people's faces, rendering the identity of the personal unrecognizable,[11] and this can be misleading.

If deepfake technology is used to portray a person favourably or unfavourably, this can obviously affect the perception that other people may have of this person. Potentially, this could threaten democratic elections, if people are portrayed saying things that significantly differ from their actual viewpoints.[12] Misleading messages can also be used to incite people to criminal behaviour or even acts of terrorism.

Another type of deception using deepfakes is the possibility to create pornographic images of celebrities (Fig. 11.2).[13] This technology 'undresses' people,

[7] Custers et al. 2019, pp. 728–745.

[8] Jingguo et al. 2012, pp. 345–362.

[9] Europol 2020.

[10] See, for instance, www.thispersondoesnotexist.com.

[11] Source: Facebook.

[12] See https://www.youtube.com/watch?v=T76bK2t2r8g.

[13] Popova 2020.

Fig. 11.1 Original and deepfake image of a woman [*Source* Facebook][14]

Fig. 11.2 Deepfake technology can be used for creating pornographic images of celebrities [*Source* https://www.ethicsforge.cc/deepfake-the-age-of-disinformation/][15]

by merging footage of celebrities with pornographic images. Actresses like Emma Watson, Natalie Portman and Gal Gadot were victims of this practice.[16] Also people who are not famous are increasingly victimised by pornographic deepfakes. This kind of footage can severely ruin people's reputations, (usually a tort, but potentially also constituting criminal acts like insult, libel or slander) deeply affecting their lives, particularly if the images become widely disseminated online.[17]

Another highly controversial type of deepfakes is the creation of virtual child pornography. Although it could be argued that this does not involve child abuse, it could lead to this. It is for this reason that virtual child pornography is a criminal

[14] Source: Facebook.

[15] Source: https://www.ethicsforge.cc/deepfake-the-age-of-disinformation/.

[16] Lee 2018.

[17] See https://www.elle.com/uk/life-and-culture/a30748079/deepfake-porn/.

offence in many countries, via the implementation of the Convention on Cybercrime[18] and EU Directive 2011/92/EU combating the sexual abuse and sexual exploitation of children and child pornography.[19]

Yet another type of deepfakes is the creation of (images of) new or different persons. The current technology allows for generating highly realistic footage of existing and non-existing persons. In the former category, deceased people can be brought back to life and incorporated in present-day images. This can go well beyond entertainment,[20] as was clearly shown in 2019, when the president of Gabon addressed his country in a deepfake video.[21] This was after months without public appearance due to hospitalisation abroad. The video led to all kinds of speculations and, shortly after, a coup attempt. Footage of non-existing persons can, in the long term, raise even more confusion. Persons only known from the screens, may very well not exist at all. When deepfakes are applied as actors, the risk may be unemployment of human actors, but when deepfakes are applied as politicians, it may become untraceable who really has the power in a country.

11.2.1.3 *Algorithmic Profiling*

Next to these mostly visual applications of AI technology, also types of AI using other types of data exist. Like in other sectors of society, also criminals make use of *profiling*,[22] a technique that can help identify characteristics and preferences of people. Criminals can use this to convince victims, as described above, but also to select which individuals and groups of people may be easy or wealthy targets.

One thing that differentiates profiling from A/B testing is that profiling requires the processing of personal data, for instance, via cookies and other online trackers. Criminals can select potential victims on the basis of preferences that internet users reveal, either explicitly or implicitly, for instance, through reading and clicking behaviour. Also money mules for the laundering of criminal profits can be recruited in this way.[23]

Other types of cybercrime that make use of these approaches are CEO fraud and WhatsApp fraud. Both exist in different varieties, but CEO fraud usually boils down to sending an order to a company's financial department, on behalf of the CEO (or perhaps the CFO), to transfer money. WhatsApp fraud usually boils down to a criminal imposing a friend or family member in urgent need of money. For both types

---

[18] Convention on Cybercrime, Budapest, 23 November 2001, https://www.coe.int/en/web/conventions/full-list/-/conventions/rms/0900001680081561.

[19] https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32011L0093&from=NL.

[20] Like the painter Dali who is brought to life by the use of deepfake technology in the Dali Museum, see Lee 2019.

[21] Cahlan 2020.

[22] Custers 2013.

[23] Custers et al. 2020, pp. 121–152.

of cybercrime, criminals first need to collect personal data on their victims and on the person whom they like to impose.

11.2.2 Sanctions and Justice-Related Programmes

11.2.2.1 Evidence-Based Sanctioning

One of the most important goals of sanctions and other justice-related programs is specific prevention (or specific deterrence), i.e., preventing the perpetrator from committing another crime in the future.[24] On the basis of large amounts of data and with the use of automated analyses, empirical research can be done on which interventions yield the best results in terms of reducing recidivism. This research problem can be modelled in the same way as a doctor treating a patient: on the basis of the disease or condition (and increasingly also the characteristics of the patient),[25] the doctor determines the best medication, therapy or treatment. Similarly, courts, judges and mediators can 'administer' interventions depending on the characteristics of perpetrators (such as the crime and the situation in which the crime was committed, but also personality traits of the perpetrator and the victim).[26] All this can be included in assessing which intervention is the most effective in terms of reducing recidivism (i.e., recidivism as classifier in the models). Potential 'treatments' and 'therapies' include the type of sanction (imprisonment, community service, or a fine), conditional or unconditional sentences, probation, parole, and the eligibility and expected effectiveness of justice-related programs (such as training and education programs, for instance, focused on improving cognitive or social skills, or dealing with aggressive behaviour or addictions).

This evidence-based algorithmic profiling approach in sanctioning can be applied on a group level (what works best for specific categories of people) or at an individual level (what works best in a specific case). At both levels applications already exist in several countries. In the United States, the National Institute of Justice publishes evaluation research on its website Crime Solutions.[27] For each justice-related program it is indicated whether it is effective or not. In the Netherlands, the government publishes data on recidivism at an aggregated level via a system called REPRIS.[28] On the basis of these and other evaluation research results, an expert committee examines the programs on their quality and effectiveness and then decides on officially recognizing them.[29] A lot of research in this field is still traditional empirical

[24] As opposed to general prevention (or general deterrence), which aims to deter others than the perpetrator, mostly by setting an example to others when imposing a sanction in a specific case.

[25] This is referred to as personalized medicine.

[26] Cf. Weijer and Leukfeldt 2017, pp. 407–412.

[27] https://crimesolutions.ojp.gov.

[28] https://data.overheid.nl/dataset/repris.

[29] https://www.justitieleinterventies.nl.

research, but analyses are increasingly automated to include larger amounts of data in these evaluations. Obviously, this may entail some risks, which will be discussed below.

11.2.2.2 Instruments for Risk Assessments

Also at an individual level, this approach has added value, particularly for instruments for risk assessments. Instruments for risk assessments are commonly used in criminal law, for instance, when courts and judges are considering probation or parole. In several of the United States, the system COMPAS is used to assess recidivism risks.[30] Courts heavily weigh these models (or rather the results they spit out) in their decisions. In the Netherlands, the probation services use a system called RISC. Part of that is OXREC, an actuarial risk assessment tool that can be used to predict statistical risks.[31] These models increasingly play a role in the work of probation services and the decisions of courts.

The use of such models offers several benefits: assessments can be done in more structured and objective ways. Subjective assessors can be prone to human failure or can be influenced by bias and prejudice. If the models are self-learning, they can also recognize and incorporate new trends and developments. This obviously can also increase efficiency and reduce costs. However, there is also criticism with regard to this way of working, because the instruments do not seem to outperform assessments by human experts and there are risks involved, such as bias that can lead to discrimination.[32] In the United States, COMPAS seemed to systematically assign higher recidivism risks to Afro-Americans.[33] It is often argued that these models do not process any ethnicity data and, therefore, cannot be discriminating.[34] However, characteristics like ethnicity can easily be predicted and are therefore often reconstructed by self-learning technologies, without being visible for users.[35] Caution is advised.

11.2.3 Legal Questions

From the above subsections, it becomes clear that AI entails a substantial change in the criminal law domain. In three categories of legal questions can be distilled for

[30] https://assets.documentcloud.org/documents/2840784/Practitioner-s-Guide-to-COMPAS-Core.pdf.

[31] https://oxrisk.com/oxrec-nl-2-backup/.

[32] Van Dijck 2020.

[33] Angwin et al. 2016.

[34] Maas et al. 2020, pp. 2055–2059.

[35] Cf. Kamiran et al. 2013.

substantive criminal law. The first category concerns questions regarding the interpretation of existing law and legislation. This concerns questions on whether particular actions are covered by specific provisions in criminal codes. For instance, it can be investigated which technologies qualify as a 'computer system' in the Convention on Cybercrime. The second category concerns questions regarding which actions or behaviour should be considered criminal, even though it may not (yet) be criminal according to the provisions in criminal codes. For instance, it may be argued that several types of deepfake technology should perhaps be prohibited by criminal law. The third category concerns questions regarding the use of data. These are questions regarding the extent to which data can be collected and processed, for instance, in the risk assessments discussed above, or questions regarding proportionality, to protect the interests of others involved, such as intellectual property of profiles and other knowledge, privacy, and equal treatment (not only of suspects, but also of non-suspects in control groups).

11.3 AI and Procedural Criminal Law

11.3.1 Criminal Investigation

Law enforcement agencies and public prosecution services can also use AI in different ways. In criminal investigation and prosecution, AI can support or even replace some parts of the work. This section will provide examples of both developments. In this section, predictive policing and cyber agent technology are discussed as examples of AI in criminal investigation and prosecution.

11.3.1.1 Predictive Policing

With the use of large amounts of data and sophisticated data analytics, trends and developments in crime can be disclosed. These technologies can also be used to predict crime, including the locations where crime is likely to take place, who perhaps will be a criminal or a victim of crime, and how criminal networks and criminal careers may develop. This is referred to as *predictive policing*.[36]

A typical example here are so-called *crime heat maps* (Fig. 11.3), in which crime rates are visualised on maps of metropolitan areas. On such maps, neighbourhoods with high crime rates ('hot spots') can easily be recognized. With the help of AI, not only static maps with snapshots can be created, but also dynamic, real-time maps can be generated. Looking back in time then becomes possible, but also looking

[36] Ferguson 2019; Schuilenburg 2016, pp. 931–936.

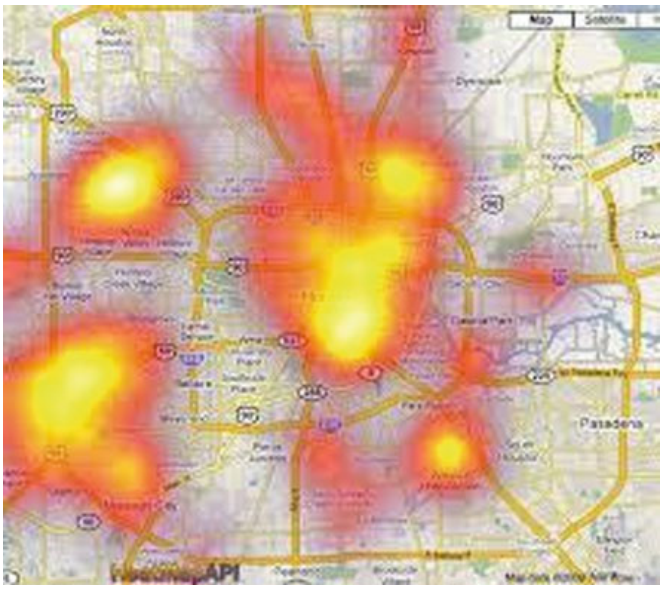

Fig. 11.3 *Crime heat maps* show crime rates for each neighbourhood. With the use of AI, real-time and prospective maps can be generated [*Source* https://spotcrime.wordpress.com/2009/07/20/houston-crime-map-new-data-and-shooting-heat-map/][37]

forward in time, by incorporating prediction models in the maps. This makes such maps useful when planning surveillance and developing policing strategies.[38]

Predictive policing can be based on location, but also on persons. With the use of profiling strategies described above, predictions can be made regarding who may commit a crime. This may be relevant for recidivism, but also for first offenders. On the basis of personal and situational characteristics, it can be predicted who constitutes a high risk to become a criminal.[39] AI related technologies can discover novel, unexpected patterns in this area and provide real time information, for instance, by also including social media data in the models. Real time information allows law enforcement to intervene on the spot, when the probability of catching a criminal is the highest. Although this approach may offer benefits in terms of efficiency and effectiveness, it should be used with caution, though: there may be crime displacement,[40] there may be disparate impact,[41] and there may be tunnel vision, with false positive and false negative rates resulting from limited reliability.[42]

[37] Source: https://spotcrime.wordpress.com/2009/07/20/houston-crime-map-new-data-and-shooting-heat-map/.

[38] Weisburd and Telep 2014, pp. 200–220.

[39] Kleemans and De Poot 2008, pp. 69–98.

[40] Weisburd et al. 2006, pp. 549–592.

[41] Barocas and Selbst 2016.

[42] Custers 2003, pp. 290–295.

11.3.1.2 *Cyber Agent Technology*

Crime rates have been steadily decreasing for many years in Western countries, but this does not seem to apply to cybercrime. In fact, for cybercrime, there seems to be an increase. That may not be surprising, since for cybercriminals the chances of being caught are low and the profits can be very high compared to offline crime. Also traditional types of crime, such as organised crime groups trafficking and trading drugs, have gone online, via online marketplaces on the darkweb (the part of the internet that has not been indexed by search engines and is only accessible with special software). One of the first illegal market places was Silk Road, established in 2011 and taken down by the FBI in 2013, where illegal substances and weapons were traded and even the services of hitmen could be purchased. After Silk Road was taken down, other websites followed, including Silk Road 2.0 (in 2014), Evolution (in 2015), AlphaBay (in 2015), Hansa (in 2017), Outlaw (in 2017), Digital Shadows (in 2018), Dream Market (in 2019), DeepDotWeb (in 2019) and Darkmarket (in 2021).[43]

It can be complicated and time-consuming for law enforcement agencies to monitor activities on these online marketplaces. For instance, access to these marketplaces requires carefully building a reputation, as the criminals on the platforms are very reluctant to allow access to new people. For law enforcement agencies it may also be required to use extensive criminal investigation competences, including the use of systemic surveillance, working undercover, secretly recording private conversations, and infiltrating in criminal organisation. Obviously, such police competences may differ per jurisdiction. Usually these competences can only be applied after a court has approved this. When applied, law enforcement agencies should be very careful not to use these competences in ways that may be seen as entrapment, as this may render any evidence collected useless in courts.

Due to the invasive and precarious nature of criminal investigations on darkweb marketplaces, it may be helpful to deploy AI. This can be done with cyber agent technology, i.e., technology that supports cyber agents (online actors). This technology can have a certain degree of autonomy and act according to the circumstances.[44] This is intelligent software that can interact with others and act without human intervention.[45] With the use of this technology, many more interactions with actors on darkweb forums can be maintained than human law enforcement officers could take care of.

One of the most concrete applications in this area is a chatbot (an automated interlocutor) called Sweetie (Fig. 11.4).[46] The chatbot is designed to look like a 10-year-old girl from the Philippines and can have conversations online with people that show sexual interests in children. The goal obviously is to track and identify

[43] For more background, see also Mirea et al. 2019, pp. 102–118.

[44] Schermer 2007.

[45] Nwana 1996, 205–244; Luck et al. 2004, 203–252.

[46] https://www.terredeshommes.nl/programmas/sweetie-20-webcamseks-met-kinderen-de-wereld-uit. See also van der Wal 2016.

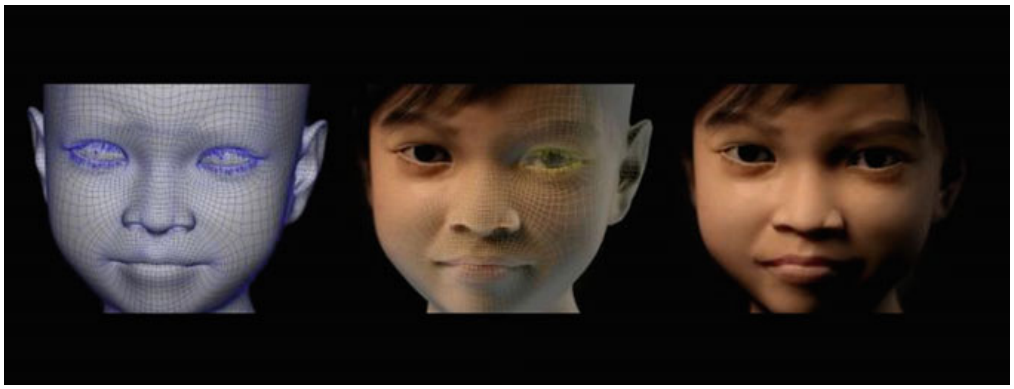

Fig. 11.4 Sweetie 2.0 is cyber agent technology that can contribute to online criminal investigation [*Source* https://www.universiteitleiden.nl/en/news/2019/07/sweetie-2.0-using-artificial-intelligence-to-fight-webcam-child-sex-tourism][47]

paedophiles and to prosecute or rebuke them. This technology can also be used by secret services and intelligence agencies in the interest of national security.[48]

The AI technology can only be used if it is sufficiently advanced, i.e., if it can pass the Turing test,[49] in which people do not realise that they are communicating with AI. In the case of Sweetie 2.0, the Turing test was not an issue: approximately 20.000 men from 71 countries reached out to her, believing she was a real child.[50] Furthermore, the technology cannot provoke illegal behaviour and should not learn and adopt criminal behaviour itself.[51] That criterion is significantly more complicated to meet: in many jurisdictions, the technology can be qualified as entrapment under criminal procedure codes.[52] Another issue was that convictions for child abuse proved to be difficult in several jurisdictions, as there was no real abuse (it is impossible to sexually abuse software).[53] And even if intentions to commit child abuse constituted a criminal act, it could be hard to prove, since Sweetie is no real child. Nevertheless, the technologies led to convictions in Australia, Belgium and the UK.[54]

---

[47] Source: https://www.universiteitleiden.nl/en/news/2019/07/sweetie-2.0-using-artificial-intelligence-to-fight-webcam-child-sex-tourism.

[48] Custers 2017.

[49] Turing 1950, pp. 433–460.

[50] http://www.dawn.com/news/1054244.

[51] Like Microsoft's chatbot Tay, which started using racist language a few hours after it was released, see: Mason 2016.

[52] van der Hof et al. 2019.

[53] Schermer et al. 2019, pp. 1–94.

[54] https://nl.wikipedia.org/wiki/Sweetie_%28virtueel_personage%29.

11.3.2 Evidence

When collecting and assessing forensic evidence, AI can play a role in different ways. This section discusses searching large amounts of data that are collected during seizures, assessing evidence, and building scenarios for reconstructing crimes.

11.3.2.1 Searching Large Amounts of Data after Seizure

In specific situations and under certain conditions (usually including a court warrant), law enforcement officers can seize digital storage devices for further searching.[55] Law enforcement can let forensics experts search the devices, including smartphones, tablets, laptops, and USB keys, for evidence. Apart from issues with damaged devices or encryption, a major problem in digital forensics often is the tremendous volume of the data on these devices. Oftentimes, only small pieces of information turn out to be relevant as evidence, for instance, to complete parts of an irrefutable narrative. In fact, these are needle-in-the-haystack kind of problems and AI can be useful in addressing these problems.[56]

In the Netherlands, the National Forensics Institute developed a tool for this, called Hansken.[57] This system, an example of big data analytics, can process large amounts of data from different sources and in different formats (such as text, video, audio, etc.), including storage, indexation and making the data searchable. The labelling of data is automated. The searchability of the seized data increases the effectiveness of criminal investigations, since relevant data is overlooked less often.[58] Also, Hansken delivers very fast results, which is a major benefit in criminal investigations, in which the first 48 hours are often the most crucial and decisive, both with regard to identifying, tracing, and finding suspects and with regard to collecting forensic evidence.

11.3.2.2 Assessing Evidence

Criminal evidence exists in different types and sizes. Technical evidence, such as DNA, fingerprints, ballistics reports, always come with margins or error. In turn, this can lead to false positives and false negatives, for instance, when matching DNA found at a crime scene with DNA profiles in databases. Also, the DNA secured by forensic experts at a crime scene is a mixture of traces of DNA. With the help of AI,

---

[55] Or data can be intercepted, see Custers 2008, pp. 94–100.

[56] Hoelz et al. 2009, pp. 883–888.

[57] https://www.forensischinstituut.nl/forensisch-onderzoek/hansken. See also van Beek et al. 2015, pp. 20–38.

[58] Sunde and Dror 2021.

so-called *probabilistic genotyping* is possible, which can be used to assess whether someone's DNA really is in these mixed traces of DNA found at the crime scene.[59]

When assessing the reliability, the focus is often on the probability of a match (for instance, a 95% likelihood), but also the reliability of this probability is important (for instance, with an error margin of 3%, a likelihood in the range of 92–98%). In case of an error margin of 2%, the probability of a match was determined much more precisely than in case of an error margin of 12%. With the help of very large numbers of data and self-learning systems, the reliability of the matches can be assessed more precisely, reducing the error margins. In this way, the reliability of the evidence can be quantified much more precisely, with smaller error margins, resulting in increased reliability of the forensic evidence.[60]

The use of AI in forensics does entail some risks. Obviously, the data may contain errors and humans are not so great at understanding risks, which may result in errors in judgements. Also, the focus may shift from narratives to numbers, and from legal experts to technological experts, which a defendant in court may find harder to challenge. A potential problem with highly specialised expertises in forensics is that there may be only a very limited number of experts (which often know each other), entailing risks of tunnel vision. In court cases in which different kinds of highly sophisticated forensics are introduced, an issue may be that no expert is able to oversee all aspects of the case. Obviously, this oversight is the responsibility of the judges in the court, but as legal experts they may not be familiar with all intricacies of the forensic technologies used. These are well-known challenges, which may further increase with the use of AI in evidence.

11.3.2.3 Building Scenarios When Reconstructing Crimes

In behavioural psychology it is well-known that humans perform poorly when assessing probabilities and risks: often a narrative is more convincing than statistics, mostly because trough evolution humans have learned to quickly pick up any causal relationships.[61] Humans apparently are much better in assessing probabilities when presented with different scenarios. AI can contribute to constructing various scenarios that can be compared and weighed in courts.[62] This can be done by attaching different weight to the available evidence per scenario. The different scenarios can also be visualised, including the extent to which they are supported by the available evidence (see Fig. 11.5). In this way, it becomes clear which parts of a particular scenario need further substantiation and additional or more detailed evidence.

Bayesian statistics play an important role in this, to express conditional probabilities. A conditional probability is a probability that includes other evidence or, more precisely, the probability of the extent to which other evidence supports a

[59] Kwong 2017, pp. 275–301.

[60] Kwan et al. 2008.

[61] Kahnemann 2012.

[62] Bex et al. 2016, pp. 22–29; Schraagen et al. 2018.

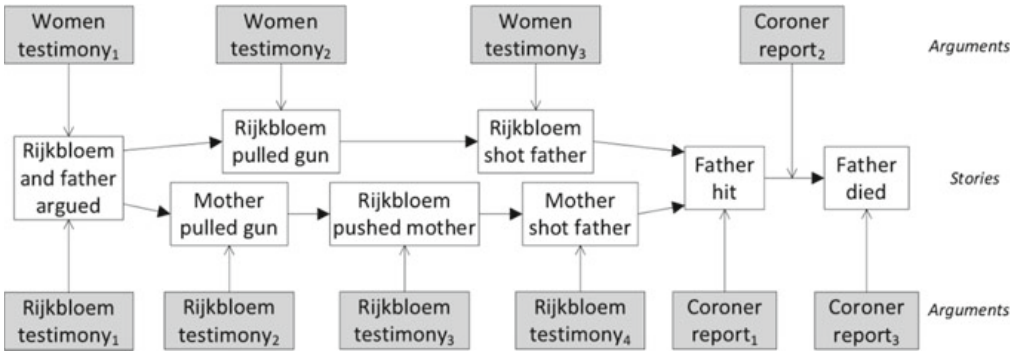

Fig. 11.5 AI technology can help construct different scenarios by varying the weight of different pieces of evidence [*Source* Bex 2015].[63]

proposition. This can be helpful in reducing the numbers of potential suspects or scenarios in a case. With the help of AI, conditional probabilities can be calculated automatically for different combinations of conditions. In other words, the AI cannot only contribute to comparing and weighing scenarios, but also to developing novel, perhaps unexpected scenarios.

11.3.3 Legal Questions

The examples in the area of procedural criminal law presented in this section may raise several ethical and legal questions. Incorrect and incomplete data, the choice of instruments for data analysis, and the interpretation of discovered profile scan all lead to limited reliability of the conclusions that are drawn. As a result of this, prejudice and discrimination may sneak into the process of criminal investigation, prosecution and sentencing. This raises ethical and legal questions with regard to substantive justice (for instance, what are suitable sentences for new types of AI enabled crime) and procedural justice (for instance, with regard to the right to a fair trial). Since AI is complex and its workings can be non-transparent or hard to explain, it may be difficult for suspects to defend themselves against this. If decisions in criminal law procedures increasingly rely on the results of AI, this could lead to situations similar to those in Kafka's novel The Trial,[64] in which suspects do not know what they are accused of, where the accusations come from, and on which information (data, analysis, conclusions) these accusations are based.

More traditional, legal positivist questions relate to the scope of the competences of law enforcement agencies. Questions include how far police powers extend in this new context of AI, how entrapment can be prevented, and how it can be guaranteed that self-learning AI will not show criminal behaviour itself after operating for some

[63] Source: Bex 2015.

[64] Kafka 2015.

time in a criminal context. Apart from interpreting the extent of existing criminal investigation competences, an important question is whether these competences are actually sufficient for criminal investigations in this rapidly developing context. This is not to argue in favour of creating more police competences, but to argue research is needed on what is perhaps missing or where existing competences can be amended to fill any gaps.

Another issue is the regulation of data analyses in criminal law. It is striking that collecting data is strictly regulated in criminal law (including data protection law), but the use of data analyses is hardly regulated.[65] In other words, once data has been collected and aggregated, law enforcement agencies and public prosecution services have a large degree for freedom to subject the data to all kinds of analyses. Regulating this could contribute to better legal protection of all actors in criminal procedures (not only suspects), for instance, via more transparency and participation. This could also increase legal certainty.

11.4 Conclusions

The goal of this chapter was to provide a concise overview of different AI developments in criminal law. The examples in this chapter illustrate that AI is increasingly used by criminals, but also by law enforcement agencies and public prosecutions services. It can be argued that the cat-and-mouse game between them has moved on to a new stage with the introduction of AI.[66] In order to keep up with developments, law enforcement agencies, public prosecution services and courts will need to invest heavily in knowledge and expertise during the next coming years.

With regard to substantive criminal law, further research is needed on the interpretation and scope of provisions in criminal codes and on whether new provisions need to be included in criminal codes in the near future to ensure that particular undesirable behaviour enabled by AI becomes punishable. With regard to procedural criminal law, further research is needed on the scope of existing criminal investigation competences, on potential modifications in these competences, and on how to properly balance criminal investigation competences and fundamental rights. The use of AI can offer many benefits in criminal investigation, but only if prejudice, discrimination, and other risks are avoided or mitigated. Regulating data analysis in criminal investigations, which is currently virtually absent, could contribute to this.

[65] Custers and Stevens 2021.

[66] Cf. similarities of other technologies introduced previously in the security domain: Teeuw et al. 2008.

References

Angwin J, Larson J, Mattu S, Kirchner L (2016) Machine Bias. ProPublica, 23 May 2016. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

Barocas S, Selbst AD (2016) Big Data's Disparate Impact. 104 California Law Review 671.

Bex FJ (2015) An integrated theory of causal scenarios and evidential arguments. In: Proceedings of the 15th International Conference on Artificial Intelligence and Law (ICAIL 2015), 13–22, ACM Press, New York.

Bex FJ, Testerink B, Peters J (2016) AI for Online Criminal Complaints: From Natural Dialogues to Structured Scenarios. ECAI 2016 workshop on Artificial Intelligence for Justice (AI4J), The Hague, August 2016, pp. 22–29.

Cahlan S (2020) How misinformation helped spark an attempted coup in Gabon. The Washington Post, 13 February 2020, https://www.washingtonpost.com/politics/2020/02/13/how-sick-president-suspect-video-helped-sparked-an-attempted-coup-gabon/.

Calders T, Custers BHM (2013) What is data mining and how does it work? In: Custers BHM, Calders T, Schermer B, Zarsky T (eds) Discrimination and Privacy in the Information Society. nr.3. Springer, Heidelberg

Calo R (2017) Artificial Intelligence Policy: A Primer and Roadmap: https://ssrn.com/abstract=3015350.

Custers BHM (2003) Effects of Unreliable Group Profiling by Means of Data Mining. In: Grieser G, Tanaka Y, Yamamoto A (eds) Lecture Notes in Artificial Intelligence. Proceedings of the 6th International Conference on Discovery Science (DS 2003) Sapporo, Japan. Springer-Verlag, Berlin/Heidelberg/New York, Vol. 2843, pp. 290–295.

Custers BHM (2008) Tapping and Data Retention in Ultrafast Communication Networks. Journal of International Commercial Law and Technology, Vol. 3, Issue 2, 2008, pp. 94–100.

Custers BHM (2013) Data Dilemmas in the Information Society. In: Custers BHM, Calders T, Schermer B, Zarsky T (eds) Discrimination and Privacy in the Information Society. Springer, Heidelberg.

Custers BHM (2017) Cyber agent technology en de Wet op de Inlichtingen- en Veiligheidsdiensten (WIV). Universiteit Leiden, Leiden, 30 September 2017.

Custers BHM (2018) Methods of data research for law. In: Mak V, Tjong Tjin Tai E, Berlee A (eds) Research Handbook in Data Science and Law. Edward Elgar, Cheltenham, pp. 355–377

Custers BHM, Pool, R, Cornelisse R (2019) Banking Malware and the Laundering of its Profits. European Journal of Criminology, Vol. 16, nr. 6, pp. 728–745. https://doi.org/10.1177/1477370818788007.

Custers BHM, Oerlemans JJ, Pool R (2020) Laundering the Profits of Ransomware: Money Laundering Methods for Vouchers and Cryptocurrencies. European Journal of Crime, Criminal Law and Criminal Justice, 28 (2020), pp. 121–152.

Custers BHM, Stevens L (2021) The Use of Data as Evidence in Dutch Criminal Courts. European Journal of Crime, Criminal Law and Criminal Justice, Vol. 29, No, 1.

Europol (2020) The Internet Organised Crime Threat Assessment (IOCTA) 2021. Europol, The Hague. https://www.europol.europa.eu/activities-services/main-reports/internet-organised-crime-threat-assessment-iocta-2020.

Ferguson AG (2019) Predictive Policing Theory. In: Rice Lave T, Miller EJ (eds) The Cambridge Handbook of Policing in the United States. Cambridge University Press.

Gallo A (2017) A Refresher on A/B Testing. Harvard Business Review, 28 June 2017, https://hbr.org/2017/06/a-refresher-on-ab-testing.

Harper J (2021) Federal AI Spending to Top $6 Billion. National Defense Magazine, 10 February 2021, https://www.nationaldefensemagazine.org/articles/2021/2/10/federal-ai-spending-to-top-$6-billion.

Hoelz B, Ralha C, Geeverghese R (2009) Artificial intelligence applied to computer forensics. Proceedings of the ACM Symposium on Applied Computing. Honolulu, 9-12 March 2009, pp. 883–888.

Jingguo W, Herath T, Rui C, Vishwanath A, Rao HR (2012) Phishing susceptibility: An investigation into the processing of a targeted spear phishing email. IEEE Trans. Prof. Commun., vol. 55, no. 4, pp. 345–362.
Jobin A, Ienca M, Vayena E (2019) The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), pp. 389–399.
Kafka F (2015) The Trial. Penguin Books, London.
Kahnemann D (2012) Thinking, fast and slow. Penguin Books, New York.
Kamiran F, Calders T, Pechenizkiy M (2013) Techniques for discrimination-free predictive models. In: Custers BHM et al. (eds) Discrimination and Privacy in the Information Society. Springer, Heidelberg.
Kleemans ER, De Poot CJ (2008) Criminal Careers in Organized Crime and Social Opportunity Structure. European Journal of Criminology. Vol. 5 Nr. 1, pp. 69–98.
Kohavi R, Thomke S (2017) The Surprising Power of Online Experiments. Harvard Business Review, September 2017, pp. 74–82.
Kwan M, Chow KP, Law F, Lai P (2008) Reasoning About Evidence Using Bayesian Networks. In: Ray I, Shenoi S (eds) Advances in Digital Forensics IV, IFIP — The International Federation for Information Processing. Springer, Heidelberg.
Kwong K (2017) The Algorithm Says You Did it: The Use of Black Box Algorithms to Analyse Complex DNA Evidence. Harvard Journal of Law & Technology, Vol., 31, Nr. 1, pp. 275–301.
Lee D (2018) Deepfakes porn has serious consequences. BBC News, 3 February 2018. https://www.bbc.com/news/technology-42912529.
Lee D (2019) Deepfake Salvador Dali takes selfies with museum visitors: it's surreal, all right. The Verge, 10 May 2019. https://www.theverge.com/2019/5/10/18540953/salvador-dali-lives-deepfake-museum.
Luck M, McBurney P, Preist C (2004) A Manifesto for Agent Technology: Towards Next Generation Computing, Autonomous Agents and Multi-Agent Systems, 9, 203–252.
Maas M, Legters E, Fazel S (2020) Professional en risicotaxatie-instrument hand in hand: hoe de reclassering risico's inschat. NJB afl. 28, pp. 2055–2059.
Mason P (2016) Racist hijacking of Microsoft's chatbot shows how the internet teems with hate. The Guardian, 29 March 2016.
Mirea M, Wang V, Jung J (2019) The not so dark side of the darknet: a qualitative study. Security Journal, 32, pp. 102–118.
Nwana HS (1996) Software Agents: An Overview. Knowledge Engineering Review. 21 (3): 205–244.
Popova M (2020) Reading out of context: pornographic deepfakes, celebrity and intimacy. Porn Studies, 7:4, 367–381, DOI: https://doi.org/10.1080/23268743.2019.1675090.
Rosemain M, Rose M (2018) France to spend $1.8 billion on AI to compete with US, China, 29 March 2018. Reuters, https://www.reuters.com/article/us-france-tech/france-to-spend-1-8-billion-on-ai-to-compete-with-u-s-china-idUKKBN1H51XP.
Schermer BW (2007) Software agents, surveillance, and the right to privacy: a legislative framework for agent-enabled surveillance. Leiden University Press, Leiden.
Schermer BW, Georgieva I, van der Hof S, Koops BJ (2019) Legal aspects of Sweetie 2.0. In: van der Hof S, Georgieva I, Schermer BW, Koops BJ (eds) Sweetie 2.0. Using artificial intelligence to fight webcam child sex tourism. Information Technology & Law Series nr. 31. T.M.C. Asser Press, The Hague, pp. 1–94.
Schraagen M, Testerink B, Odekerken D, Bex F (2018) Argumentation-driven information extraction for online crime reports. CKIM 2018 International Workshop on Legal Data Analysis and Mining (LeDAM 2018), CEUR Workshop Proceedings.
Schuilenburg M (2016) Predictive policing: de opkomst van gedachtepolitie? Ars Aequi, December 2016, pp. 931–936.
Sunde N, Dror I (2021) A Hierarchy of Expert Performance (HEP) applied to Digital Forensics: Reliability and Biasability in Digital Forensics Decision Making. Forensic Science International: Digital Investigation, Vol. 37. https://doi.org/10.1016/j.fsidi.2021.301175.

Teeuw WB, Vedder AH, Custers BHM, Dorbeck-Jung BR, Faber ECC, Iacob SM, Koops B-J, Leenes RE, de Poot HJG, Rip A, Vudisa JN (2008) Security Applications for Converging Technologies: Impact on the constitutional state and the legal order. O&B 269. WODC, The Hague.

Turing A (1950) Computing machinery and intelligence. Mind 59, pp. 433–460.

van Beek HMA, van Eijk EJ, van Baar RB, Ugen M, Bodde JNC, Siemelink AJ (2015) Digital Forensics as a Service: Game On. Digital Investigation, Vol. 15, pp. 20–38.

van der Hof S, Georgieva I, Schermer BW, Koops BJ (2019) Sweetie 2.0. Using artificial intelligence to fight webcam child sex tourism. Information Technology & Law Series nr. 31. T.M.C. Asser Press, The Hague.

Van Dijck G (2020) Algoritmische risicotaxatie van recidive: over de Oxford Risk of Recidivism tool (OXREC), ongelijke behandeling en discriminatie in strafzaken. NJB 2020/1558.

van der Wal C (2016) Sweetie 2.0: nieuw virtueel meisje gaat op pedojacht, Algemeen Dagblad, 13 February 2016. https://www.ad.nl/binnenland/sweetie-2-0-nieuw-virtueel-meisje-gaat-op-pedojacht~ad3739ca/.

Weijer SGA, Leukfeldt ER (2017) Big Five Personality Traits of Cybercrime Victims. Cyberpsychology, Behavior, and Social Networking, 20(7), pp. 407–412.

Weisburd D, Telep CW (2014) Hot Spots Policing. Journal of Contemporary Criminal Justice, 30(2), pp. 200–220.

Weisburd D, Wyckoff LA, Ready J, Eck JE, Hinkle JC, Gajewski F (2006) Does Crime Just Move Around the Corner? A Controlled Study of Spatial Displacement and Diffusion of Crime Control Benefits. Criminology 44 (3), pp. 549–592.

Bart Custers PhD MSc LLM is (full) professor of Law and Data Science and director of *eLaw, Center for Law and Digital Technologies* at Leiden University, the Netherlands. He has a background in both law and physics and is an expert in the area of law and digital technologies, including topics like profiling, big data, privacy, discrimination, cybercrime, technology in policing and artificial intelligence. As a researcher and project manager, he acquired and executed research for the European Commission, NWO (the National Research Council in the Netherlands), the Dutch national government, local government agencies, large corporations and SMEs. Until 2016 he was the head of the research department on Crime, Law enforcement and Sanctions of the scientific research center (WODC) of the Ministry of Security and Justice in the Netherlands. Before that, he worked for the national government as a senior policy advisor for consecutive Ministers of Justice (2009–2013) and for a large consultancy firm as a senior management consultant on information strategies (2005–2009). On behalf of the Faculty of Law, he is the coordinator of the SAILS project. This project, funded by the Executive Board of Leiden University, deals with the societal and legal implications of Artificial Intelligence. Bart Custers published three books on profiling, privacy, discrimination and big data, two books on the use of drones and one book on the use of bitcoins for money laundering cybercrime profits. On a regular basis he gives lectures on profiling, privacy and big data and related topics. He has presented his work at international conferences in the United States, Canada, China, Japan, Korea, Malaysia, Thailand, the Middle East and throughout Europe. He has published his work, over a hundred publications, in scientific and professional journals and in newspapers.