ALIGNER

ALIGNER D2.2

Archetypical Scenarios and their Structure

Deliverable No.	D2.2
Work Package	WP2
Dissemination Level	PU
Author(s)	Lindsay Clutterbuck (CBRNE Ltd)
Co-Author(s)	-
Contributor(s)	Richard Warnes (CBRNE Ltd)
Due date	2022-03-31
Actual submission date	2022-11-09
Status	Final
Revision	2.0
Reviewed by (if applicable)	Ezgi Eren (KUL), Daniel Lückerath (Fraunhofer), Tommy Westman (FOI), Andrius Paškauskas, Willie Koolhof (LEAAB), Catherine Aleppo, Oliver Rose (SIEAB)

This document has been prepared in the framework of the European project ALIGNER – Artificial Intelligence Roadmap for Policing and Law Enforcement. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 101020574.

Contact:

info@aligner-h2020.eu
www.aligner-h2020.eu

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement no. 101020574.

Version history

Version	Date	Comments
1.0	2022-03-29	Originally submitted version
2.0	2022-11-09	Added clarification in which tasks the scenario narratives will be continued to be elaborated • p. 6, first paragraph • p. 28, new section 5.3 on the first ALIGNER Scenario

Executive Summary

The aim of Project ALIGNER is to bring together *"…European actors concerned with Artificial Intelligence (AI), Policing and Law Enforcement and to collectively identify and discuss needs for paving the way for a more secure Europe in which Artificial Intelligence supports LEAs while simultaneously empowering, benefiting and protecting the public."* An integral part of achieving this is the design and development of AI-related scenarios to act as focal points in the project work packages (WP) relating to AI 'Technology impact assessment' (WP3) and 'Ethics and Law' (WP4) and in addition, to contribute to each of the five iterations of the ALIGNER deliverable D5.3, entitled 'Research Roadmap for AI in support of law enforcement and policing'.

It is important that the scenarios and scenario narratives produced meet two criteria; they are credible, relevant and rooted in the current realities facing Police and Law Enforcement Agencies (P&LEA) while at the same time, they must also enable thinking that encompasses emerging and foreseeable developments in AI, its potential impacts on policing and law enforcement and more generally, on society itself.

To commence this work, two project workshops were held, one 'face to face' in November 2021 focused on the potential crime and security threat from AI and one 'on-line' in January 2022 focused on examining the capability enhancement needs of Police and Law Enforcement Agencies (P&LEA). A modified 'Grounded Theory' approach (i.e. working from the 'bottom up') was adopted to gather and analyse the input arising from the workshops. This was collated into two typologies of AI use.

Together, they enabled ALIGNER to explore the concept that the nature of AI in the context of policing and law enforcement has two dimensions; one where AI can be a crime and security threat and one where AI can be utilised by Police and Law Enforcement Agencies to increase their operational capabilities and to drive improvements to their overall effectiveness. From this, ALIGNER was able to develop a structure and process to ensure that it is possible to meet the AI scenario needs of the project. It commences with the ALIGNER 'archetypical scenario'.

The word 'archetypical' is generally defined to mean an original model or event from which all subsequent related models or events are drawn. Consequently, ALIGNER interprets an "archetypical scenario" as being an original scenario from which all other similar or related scenarios are then derived. The ALIGNER archetypical scenario consists of an imagined world in which AI is a constant crime and security threat and where AI is also regularly utilised by P&LEA.

In turn, by selecting required elements from the ALIGNER archetypical scenario and the two typologies that categorise the practical uses of AI within each of its two dimensions, scenario narratives can then be written for the two types of AI scenario, namely 'AI Threat scenarios' (based on AI as a crime and security threat) and 'AI Use-case scenarios' (based on the utilisation of AI by P&LEA).

The Archetypical Scenario and the scenario framework were derived from the workshops, where over fifty examples of AI threat were discussed at the first workshop alone. These were refined and developed to give the two typologies of AI use described above. The

first typology, where AI is or could be a crime and security threat, consists of four over-arching categories plus twelve sub-categories. The main categories are;

- AI, vehicles, robots and drones
- AI, crime and criminal activity in the digital domain
- AI, disinformation and social manipulation
- AI and on-line cybercrime

The second typology covers AI when it is utilised by P&LEA. The specific AI use categories are;

- Recognition and Identification of Individuals
- Crime and threat detection and prevention
- Data and information handling processes
- Digital forensics
- Digital domain activity
- Autonomous vehicles, robots and drones

Also included here are four categories relating AI use to the core capabilities of P&LEA

- Prevention and detection capabilities
- Reaction and response capabilities
- Investigation and prosecution capabilities
- Ancillary P&LEA capabilities

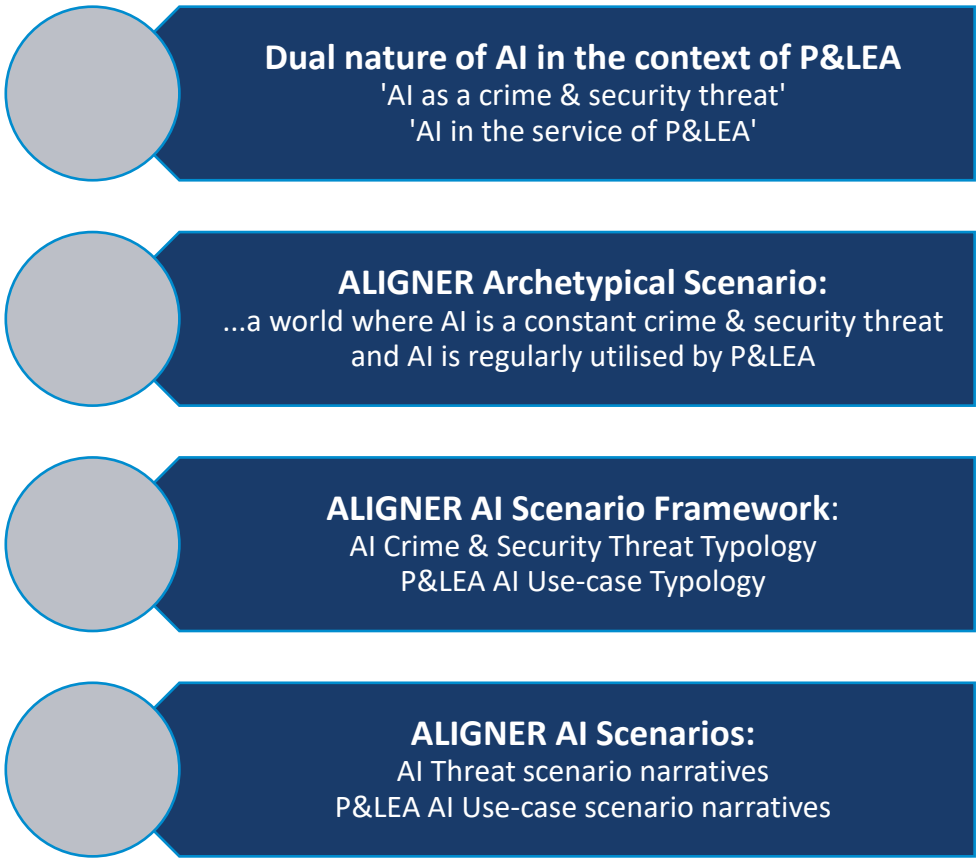
While the project requirement for ALIGNER was to create a structure of four AI "archetypical crime scenarios", it is clear that by concentrating only on the 'crime and security' dimension of AI, the other, perhaps more important dimension of AI, where AI can be utilised in the service of P&LEA, would be excluded. The approach taken by ALIGNER gives an opportunity to generate and write scenario narratives suitable for use as in 'AI Threat scenarios' and in 'AI Use-case scenarios'. While both of these originate from the Archetypical Scenario and scenario framework, the subsequent scenario narratives will need to be written differently.

AI Threat scenario narratives should, on the one hand, cover the scientific aspects of the AI technology utilised by a 'bad actor' and the modus operandi (MO), tactics, techniques and procedures (TTPs) involved in its preparation and use. Within the constraints imposed by necessary restrictions on the release of information, they could also include the scientific aspects of any AI technology utilised by P&LEA to counter the 'hostile' AI and the 'bad actors' and the MO and TTPs used by P&LEA to implement their response. Finally, the ethical, legal and social implications of any potential P&LEA response should be incorporated and explored.

On the other hand, developing relevant P&LEA AI Use-case scenarios is also beneficial. Their structure needs to include any AI threat-based element if appropriate but also, it must encompass the technological aspects of the P&LEA AI under examination, the MO and TTPs used by the P&LEA to implement the response or capability and the ethical, legal and social implications of their doing so.

The narrative for the first ALIGNER Scenario will be published in September 2022 as part of the first iteration of deliverable D5.3, *"Research Roadmap for AI in support of Law Enforcement and Policing"*. It will be an 'AI Threat scenario' and drawn from the AI Crime and Security Threat category of 'AI, disinformation and social manipulation'. The elements for inclusion in the scenario narrative are now under consideration.

Figure: Summary of the ALIGNER AI concept and overall AI scenario structure



Updates in version 2.0 from November 2022

This second version of the deliverable was produced after the ad-hoc review of the project in July 2022 upon request of the reviewers and better clarifies in which tasks the work on the scenario narratives will continue. For this purpose, a new section 5.3 was introduced.

Table of Contents

List of Abbreviations

Abbreviation	Meaning
AI	Artificial Intelligence
ALIGNER	Artificial Intelligence Roadmap for Policing and Law Enforcement
KPI	Key Performance Indicator
LEEAB	Law Enforcement Agency Advisory Board (ALIGNER)
ML	Machine Learning
MO	Modus operandi
P&LEA	Policy and Law Enforcement Agencies
SIEAB	Scientific, Industrial and Ethical Advisory Board (ALIGNER)
SO	Specific Objective
TTPs	Tactics, techniques and procedures
WP	Work package

1. Introduction

Project ALIGNER commenced in October 2021, with the overall aim of *"...[bringing] together European actors concerned with Artificial Intelligence (AI), Law Enforcement and Policing to collectively identify and discuss needs for paving the way for a more secure Europe in which Artificial Intelligence supports LEAs while simultaneously empowering, benefiting and protecting the public."*

To help achieve it, ALIGNER has set seven Specific Objectives (SO), each one supported by a number of Key Performance Indicators (KPI). Under SO 1, the KPIs relate to;

- The formation of two Advisory Boards for the project[1]
- Undertaking a series of eight workshops involving experts and practitioners from the Advisory Boards, plus invited experts and practitioners from other on-going or future EU research projects, in order to facilitate debate and exchanges between them
- The development of "*at least four archetypical crime scenarios…collectively modelling the crime pattern changes due to AI, building the basis for subsequent project activities.*"

All of these activities take place within ALIGNER work package 2 (WP2), *'Law Enforcement Agencies (LEA) and Civil Society Engagement'* and occur across the three year life-span of the project.

The aim of this deliverable (*'D2.2- Archetypical scenarios and their structure')* is to document the steps taken in the earliest stage of ALIGNER to develop *"…a systematic (scenario) description method to be employed to identify and analyse scenarios relating to needs, consequences and recommendations from a practitioners point of view."* It shows how the first two ALIGNER workshops were used as part of this to gather data and test ideas. These were then further developed into the concept that AI in the context of policing and law enforcement has a dual nature, consisting of two intertwined dimensions, one where AI is a crime and security threat that Police and Law Enforcement Agencies (P&LEA) must respond to and challenge as necessary, and the other dimension where P&LEA utilise AI in order to carry out their functions and duties.

The next steps were to develop a structure for scenarios that reflected this dichotomy. It is comprised of both an 'Archetypical Scenario' and an AI Scenario Framework that is appropriate to both dimensions of AI. From these, two types of AI Scenario can be identified; 'AI Crime and Security Threat scenarios' and 'P&LEA AI Use-case scenarios'. Once the type of AI scenario required is selected from these two, an appropriate scenario narrative can then be written for it. The scenario narrative can be designed to highlight any AI issues selected for examination (see Figure 1 below).

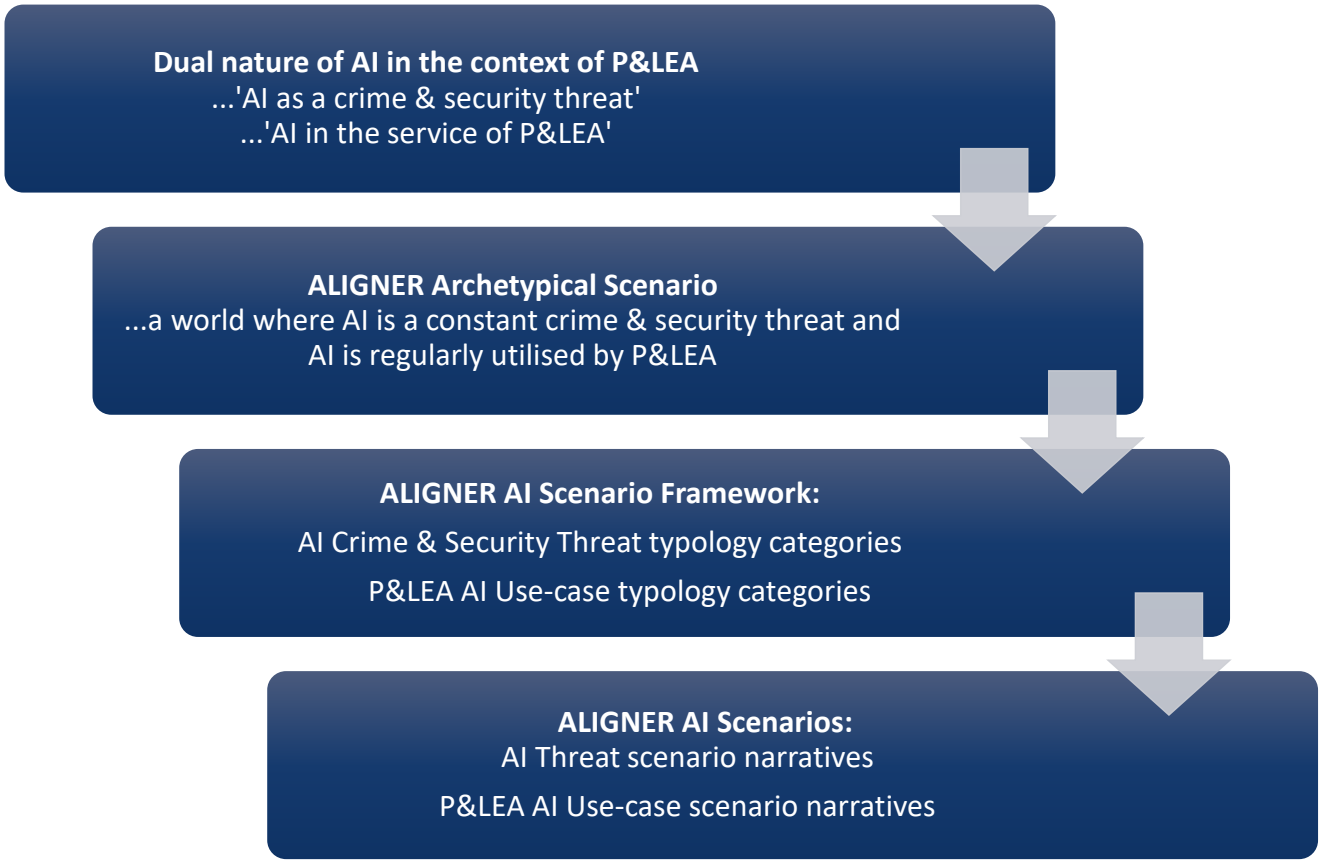
The work of ALIGNER was originally envisaged as being focused on AI "archetypical crime scenarios", where AI poses a threat to P&LEA. However, once ALIGNER began it rapidly became clear that the project had to approach AI in the context of policing and law

[1] The Law Enforcement Agency Advisory Board (LEEAB) and the Scientific, Industrial and Ethical Advisory Board (SIEAB)

enforcement in a holistic way if it was to achieve the most relevant and useful results. By concentrating only on AI as a crime threat, there would be a significant gap as the second critical dimension of AI, where AI is utilised in the service of P&LEA, would not be examined in sufficient depth.

Consequently, this deliverable shows how the concepts of an Archetypical Scenario and the Scenario Framework are also a useful mechanism to explore the AI capability enhancement needs of P&LEA. The key elements of the overall scenario structure and how they are used is shown in Figure 1 below.

Figure 1: The key elements of the ALIGNER concept of AI and the overall AI scenario structure



1.1 Relation to other Deliverables

All the elements of the scenario structure, from the Archetypical Scenario to the scenario narratives for the 'AI Threat scenarios' and 'P&LEA AI Use-case scenarios' will be used to contribute to other elements of the project, as appropriate. They will serve to inform the work within the Work Packages (WP) and that of the project collectively. These include WP5 through each iteration of its key deliverable of the *'Research Roadmap for AI in support of LE and Policing'* (D5.3)*;* WP3, *'Technology Impact Assessment'* and WP4, *'Ethics and Law'*.

WP 3 in particular will benefit from the ALIGNER scenarios, particularly in Task 3.1 - '*Continuous screening and analysis of AI* solutions', Task 3.3 -'*AI technology risk assessment*' and Task 3.4 - '*Taxonomy of AI-supported crime*'. It will also be able to draw

on the two typologies devised to collate and categorise the practical uses of AI within each of its two dimensions.

Consequently, the work arising from the scenario process described in this deliverable will also be incorporated into deliverables D3.1 - '*Impact assessment of AI Technologies for EU LEAS'*, D3.2 - *'Risk assessment of AI technologies for EU LEAS'* and D3.3 - '*Taxonomy of AI supported crime*

As well as the scenarios themselves, a number of ALIGNER Tasks will be able to draw on the data collected during and subsequent to the first two ALIGNER workshops. An outline of the two data collection grids used in the process of creating the two typologies that emerged from them is presented as Annex A and Annex B. The central role of the ALIGNER workshops in developing the AI scenario structure and process is described in Section 2.

1.2 Gender Statement

ALIGNER partners actively safeguard gender equality and are aware of gender issues in science and technology (ref. "Commission of the European Communities: Women and Science: Excellence and Innovation – Gender Equality in Science, SEC (2005) 370, available at https://data.consilium.europa.eu/doc/document/ST-7322-2005-INIT/en/pdf).

ALIGNER monitors gender equality addressing biases and constraints throughout all the stages of the project as listed in Gendered Innovations 2 (ref "European Commission: Gendered Innovation 2 How Inclusive Analysis Contributes to Research and Innovation, (2020) available at https://op.europa.eu/en/publication-detail/-/publication/33b4c99f-2e66-11eb-b27b-01aa75ed71a1/language-en).

Outreach activities, visual representations, events, modes of data gathering and analysis, and other research products related to D2.2 have been and will be gender proofed during the internal review process following the ALIGNER Gender policy (ref: ALIGNER D1.2 Project Handbook, section 8 'Gender aspects in publications and research').

2. The Role of the ALIGNER Workshops

The aim of the first ALIGNER workshop, held at KU Leuven, University of Leuven, Belgium between November 17th and 18th, 2021 was to begin to identify 'real world' information and examples that could be utilised to devise an 'archetypical scenario' and a scenario framework relevant to the topic of AI in the context of policing and law enforcement. The participative approach adopted enabled participants to contribute their own knowledge and insights of examples, case-studies and information and to explore the issues raised.

Prior to the first ALIGNER Workshop, the research output from two other AI-focused workshops run in recent years by other entities was examined. The first was held in Oxford in 2017 and entitled "*Bad Actor Risks in Artificial Intelligence*".[2] Here, selected specialists had examined AI as a security threat across three security domains, identified as Digital Security, Physical Security and Political Security. Overall, they identified nineteen *"areas of plausible concern"*. No examination of AI and its specific relationship to crime and criminality was undertaken.

The second external workshop examined was held under the auspices of University College London (UCL) in 2019 and entitled "*AI and Future Crime*".[3] It had identified twenty examples of threats that together formed an *"approximate taxonomy of criminal applications"*. Based on considerations of 'harm, criminal profit, achievability and defeatability', the workshop assigned each individual example to a relative threat level, designated as High Threat, Medium Threat or Low Threat.

At the first ALIGNER workshop, the main focus was '*AI as a Crime and Security Threat*', but a start was also made on examining the other intimately connected aspect of AI, AI in the service of Police and Law Enforcement Agencies (for full details of the methodology employed, see 'Methodology Outline' in Section 2.1 below). All of the examples from the Oxford and London workshops were examined and discussed and this in turn generated new examples and ideas of AI as a crime and security threat. From the resulting consolidated data, a typology of four categories and twelve sub-categories for AI as a crime and security threat was drawn up (see Annex A).

The second part of the first ALIGNER workshop took the same approach but this time participants were asked to concentrate on examples of where AI was, or was likely to be, utilized in the service of P&LEA, a theme that was to be taken up more fully in the next workshop.

The second ALIGNER Workshop took place on-line on January 18th and 19th, 2022 and adopted the same interactive approach as in Workshop 1 to gather from participants

---

[2] 'The Malicious use of Artificial Intelligence: Forecasting, prevention, and mitigation'
Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., Dafoe, A., Scharre, P., Zeitzoff, T., Filar, B., Anderson, H., Roff, H., Allen, G. C., Stein- hardt, J., Flynn, C., hÉigeartaigh, S. Ó., Beard, S., Belfield, H., Farquhar, S., Lyle, C., Crootof, R., Evans, O., Page, M., Bryson, J., Yampolskiy, R., & Amodei, D. *arXiv.org*, vol. cs.AI. (2018), https://arxiv.org/pdf/1802.07228.pdf?source=post_page, Accessed October 2020

[3] 'AI-enabled future crime' M.Caldwell, J.T.Andrews, T.Tanay and L.D.Griffin, Crime Science (2020), https://link.springer.com/content/pdf/10.1186/s40163-020-00123-8.pdf, Accessed October 2020

relevant knowledge and information. It comprised of interactive sessions on both days, interspersed with additional specialist presentations by members of the ALIGNER SIEAB and consortium. It continued to examine examples of the crime and security threat posed by AI but the focus was on discussing the AI and ML 'capability enhancement needs' of P&LEA today, the needs they may well have in the future and the actions required to bring AI safely and effectively into the service of P&LEA.

The new ALIGNER AI typology categories for AI as a Crime and Security Threat that arose from Workshop 1 and those for AI in the service of P&LEA were used in Workshop 2 as a framework to encourage the contribution of new information and ideas from the ALIGNER participants. How this was carried out is described in the section below.

2.1 Methodology Outline

The rationale conceived prior the commencement of ALIGNER project provided a start point for the project scenario work by recognising that "*Police and LEAs are at the forefront of dealing with the dual challenge of maximising the benefits of AI (for example, by benefitting from the advancement of more accurate facial recognition solutions) while simultaneously having to counter the tactics, techniques and procedures used to defeat the legitimate purposes of AI.*"

In order to obtain the maximum input of the experience and technical expertise of the various ALIGNER Advisory Boards, consortium members and invited participants towards the development of an archetypical scenario, the research undertaken utilised a variant of Grounded Theory and the use of data tables[4]. During discussions and presentations during ALIGNER Workshop 1 and Workshop 2, detailed notes were recorded of relevance to the dual nature of AI, namely AI as a crime and security threat and AI in the service of Police and Law Enforcement Agencies.

These workshop notes were transcribed and subjected to a variant of Grounded Theory, using a form of 'constant comparative textual analysis' where Open and Axial coding were applied to the transcripts. This entailed the fragmentation of the qualitative data to identify the key categories and sub-categories that were then 're-assembled' as a typology to identify wider, over-arching categories. The two data tables created, one for 'AI as a crime and security threat' and the other for 'AI utilisation in the service of Police and Law Enforcement Agencies' were then populated with 'buckets' of data drawn from the workshop notes to provide key examples of the identified variables (See Annex A and Annex B for the structure and categories of the two data tables. They do not show any inserted data).

The resulting analysis of the populated tables was used to inform and develop the typologies, the archetypical scenario and the scenario framework. They also assisted in

[4] Grounded Theory is a methodology that uses an inductive approach to qualitative data, enabling general conclusions to be drawn out of an assembled mass of specific data points. (See Strauss, A. & Corbin, J. (1994). Grounded Theory Methodology: An Overview. In N. Denzin & Y. Lincoln *Handbook of Qualitative Research. 1st ed.* (pp 273-284) and Warnes, R. (2009) Grounded Theory. In Ling, T. & Villalba van Dijk, L. *Performance Audit Handbook: Routes to Effective Evaluation.* Santa Monica, Rand Corporation)

defining the two types of ALIGNER AI scenario, i.e. AI Threat scenarios and P& LEA AI Use-case scenarios. Finally, the tables will be updated and referred to for the creation of scenario narratives throughout the duration of the project and hence they will contribute on a continuous basis to each iteration of the ALIGNER research roadmap.

This research approach has been developed and refined in various field research over the last decade, and has proved particularly useful in examining 'real world' operational events as a precursor to developing contingency planning scenarios.

2.2 Utilising the results from the ALIGNER Workshops

Over the course of both ALIGNER Workshops it became clear that in the context of policing and law enforcement, AI already makes an impact across two different yet intimately interlinked dimensions.

The first is where AI is utilised by 'bad actors' to enable them to carry out acts hostile or detrimental to society and individuals within it, i.e. where AI is a crime and security threat. Here, the key aspect for Police and LEA's is that they have no control over the AI, nor (initially) the 'bad actors' who seek to use it for their own illegal or illicit ends. This can create crime and security threats that Police and LEA's may have to respond to in both the physical domain (to counter the criminals and their criminal actions) and in the digital domain (to counter the AI they may be utilizing to carry out their criminal actions).

The second of these dimensions is where AI is utilised by Police and LEAs to fulfil their duties to protect society, i.e. where AI is in the service of policing and law enforcement. Here, the key aspect is that Police and LEA's are in control of the AI they are using, including its functions and targeting. Consequently, they can be proactive as to how it is directed and utilized and crucially, they are responsible and accountable for all aspects of its use and impact.

The examination at the workshops of both of these dimensions provided a number of outputs relevant to the above and which were used in the development of the ALIGNER archetypical scenario and scenario framework.

2.2.1 AI as a Crime and Security Threat

By integrating the examples and cases discussed during the ALIGNER workshops with the two sets of inputs drawn from the external workshops held in Oxford and London, it was possible to construct a typology of twelve categories relevant to the ALIGNER concept of 'AI as a Crime and Security Threat'. These new ALIGNER categories are as follows:

Figure 2: ALIGNER consolidated typology categories for 'AI as a Crime and Security Threat'

- AI enabled fraud & forgery
- AI enabled social engineering
- AI 'Deep Fakes'
- Weaponised autonomous drones
- Weaponised autonomous vehicles
- AI controlled robots
- AI disruption of AI systems
- AI data harvesting & exploitation
- AI disinformation & social manipulation
- Poisoned/biased AI data
- Exploitation of AI capabilities
- AI use in AI countermeasures

Following the approach taken by the University College of London workshop, the participants at ALIGNER Workshop 1 discussed and then agreed on a potential level of threat appropriate to each of the new ALIGNER categories. These are High Threat, Medium Threat or Low Threat. The results are shown in Figure 3 below.

Figure 3: 'AI as a Crime & Security Threat': ALIGNER typology categories ranked by their perceived threat levels

High Threat

AI disinformation and social manipulation

AI 'Deep Fakes'

AI enabled fraud and forgery

AI enabled social engineering

AI disruption of AI systems

Medium Threat

AI data harvesting and exploitation

Exploitation of AI capabilities

Low Threat

Weaponised autonomous vehicles

Weaponised autonomous drones

AI controlled robots

'Poisoned/biased AI data

AI use in AI counter measures

It is accepted that this is only a 'snapshot' of current opinion, but it could serve as a baseline to compare against if the ALIGNER participants were asked to repeat the risk level assessment during the final year of the project in 2023.

The twelve categories were also used to derive the four, over-arching categories of 'AI as a Crime and Security Threat (see Figure 4 below).

Figure 4: The four over-arching typology categories of 'AI as a Crime and Security Threat'

AI, vehicles, robots and drones

AI, crime & criminality in the digital domain

AI, disinformation and social manipulation

AI and on-line cybercrime

In turn, the over-arching categories have been taken into use as part of the ALIGNER scenario framework relating to 'AI as a Crime and Security Threat'.

2.2.2 AI in the service of Police and Law Enforcement Agencies

Both ALIGNER Workshop 1 and Workshop 2 examined the topic of 'AI utilised in the service of P&LEA' and categories were also derived for it in the same way as they were for 'AI as a Crime and Security Threat'. The over-arching categories in Figure 4 above provided the start point for the discussion. However, the examples, case studies and ideas led to a different final set of categories, as seen in Figure 5 below. The first six of them cover broad areas of P&LEA activities and capabilities. The next four are designed to encapsulate more specifically the core functions of P&LEA where AI is currently, or potentially could be, beneficially applied.

Figure 5: 'AI in the service of Police & Law Enforcement Agencies' (ALIGNER typology categories)

Recognition and Identification of Individuals

Crime and threat detection and prevention

Data and information handling processes

Digital forensics

Digital domain activity

Autonomous vehicles, robots and drones

Prevention and detection capabilities

Reaction and response capabilities

Investigation and prosecution capabilities

Ancillary P&LEA capabilities

In summary, the two ALIGNER Workshops provided input that can be utilized in a number of ways, particularly to create a scenario framework that defines the parameters of AI in the context of policing and law enforcement. It was also used and to begin the process of constructing and selecting the ALIGNER Archetypical Scenario.

3. The ALIGNER Archetypical Scenario

In order for ALIGNER to develop relevant and informative AI focused scenarios, a clear understanding is needed of what the word 'scenario' means in the context of AI, policing and law enforcement in the European Union. It is also necessary to understand what is meant within the project by an 'archetypical scenario'.

Scenarios are widely used and they can be very different depending on the purpose they are intended for. At a basic level, the word 'scenario' is often applied to a comprehensive and detailed analysis of facts relating to a specific historical event or set of circumstances. Their purpose is to answer the question 'What happened?' and thereby to understand why and how an event or incident occurred or a set of circumstances arose. Scenarios of this type are frequently used in the context of real-world operations and response planning.

While a scenario can be derived from existing knowledge, facts and realities, a scenario can also include imaginative or embellished elements that have been extrapolated from them to highlight selected aspects or possibilities. In these cases, the scenario is used to pose the question 'What if…?' and the answers arrived at relate to what is foreseeable, possible or even just imaginable. Consequently, this type of scenario is frequently used in contingency planning as a means to identify and explore potential threats and responses.

Finally, a scenario can be a simulation of possible futures or aspects of possible futures created from a synthesis of ideas, foresight and imagination. It may have very few elements relevant to past or contemporary events, or may even have none at all. The emphasis is not on determining what has happened in the past or what the implications of it might be but focus on envisioning what the future might hold. In this form, they are often used in futures methodologies to generate new ways of thinking about the issues they create and highlight.

As a project, ALIGNER needs to utilise the synergy that can be created by using all of these approaches to developing the narratives for AI scenarios. However, any AI scenario developed should, first and foremost, be credible, realistic and relevant to the current and foreseeable future requirements of policing and law enforcement. To achieve this, it must encapsulate facts derived from real events and circumstances. In addition, the scenario should also enable the exploration of potential futures.

The concept of the dual nature of AI in the context of policing and law enforcement enables scenarios to be used in two ways. First, to delineate and explore the crime and security threats posed by AI, i.e. 'AI Threat scenarios' and second, by highlighting the potential existence of 'AI Use-case scenarios' to explore circumstances and approaches where AI could be harnessed to service the needs of P&LEAs. These AI Use-case scenarios could be helpful in exploring the "*capability enhancement needs*" of police and law enforcement.

3.1 Constructing the ALIGNER Archetypical scenario

The word 'archetypical' is generally defined to mean an original model or event from which all subsequent related models or events are drawn. Consequently, ALIGNER interprets an 'archetypical scenario' to mean it is the original scenario from which all other similar or related scenarios are then derived.

It has already been described how, in the context of policing and law enforcement, AI makes an impact both when it is used by external actors to carry out acts hostile or detrimental to society and the individuals within it and when AI is utilised by Police and LEAs in order to fulfil their duty protecting society. By taking this concept and applying to it selected scenario planning techniques designed for use by law enforcement organisations, an archetypical scenario has been created.

On this basis, the archetypical scenario specifically generated for use in ALIGNER describes a world where, on the one hand, AI in one form or another is a constant crime and security threat to individuals, entities and society while on the other hand, AI is regularly utilised by P&LEA to carry out more effectively their responsibilities, duties and functions.

The methodology used to derive this imaginative 'world' was drawn from a toolkit devised by the "Imaginative Scenario Planning for Law Enforcement Organisations" project, funded by the Centre for Research and Evidence on Security Threats (CREST).[5] The scenario planning technique most relevant to ALIGNER, and how it was applied within the project, is outlined below.[6]

3.1.1 Identifying two factors as 'drivers of change'

From the outset of the two Workshops held by ALIGNER, the importance of the dual nature of AI in the context of policing and law enforcement was apparent. This insight also provided the two drivers of change required to create the archetypical scenario. The first driver chosen was "*AI as crime and security threat*" and the second one was *AI utilisation by Police and Law Enforcement Agencies*". The word 'utilisation' was carefully selected as its general meaning of 'making practical and effective use' fits well with how the core capabilities of policing and law enforcement operate.

It should be borne in mind that only having two drivers of change will impose limitations on any 'world' created. However, the two drivers selected by ALIGNER impact at the highest levels of AI use, both when it is used by bad actors in harmful ways or, when it is used by P&LEA in beneficial ways to carry out their duties and responsibilities. As a result, the overall outcome is more likely to be fundamental in its nature, relevant to the purpose required and produce the most useful results.

[5] www.crestresearch.ac.uk.

[6] 'Imaginative Scenario Planning for Security and Law Enforcement Organisations' by Professor Math Noortmann, Professor Juliette Koning, Dr Joost Vervoort and Dr Ingrid Hoofd, 29th November 2019, https://crestresearch.ac.uk/resources/imaginative-scenario-planning-toolkit/, Accessed December 2021

3.1.2 Developing 'two alternative future states'

Once the two drivers were identified, an alternative future state for each of them was created by defining their two polar opposites. For the first of these alternative future states (where the driving factor is AI as a crime and security threat), the polar opposites chosen ranged from where 'AI is a *constant* crime and security threat' to where 'AI is an *intermittent* crime and security threat'. The words 'constant' and 'intermittent' were chosen to give a more meaningful result than if AI was defined by absolute opposites e.g. if AI was either *always* a crime and security threat or AI was *never* a crime and security threat.

In the second alternative future state (where the driving factor is AI utilisation by Police and Law Enforcement Agencies), the polar opposites chosen ranged from where 'P&LEA *regularly* utilise AI' to where 'P&LEA *do not regularly* utilise AI'. Once more, the word 'regularly', rather than 'constantly' or 'routinely' was chosen to give a more meaningful result by generating a plausible and realistic description of potential P&LEA utilisation of AI, rather than using words to describe an unlikely state of affairs where P&LEA *never* used AI or *always* used AI.

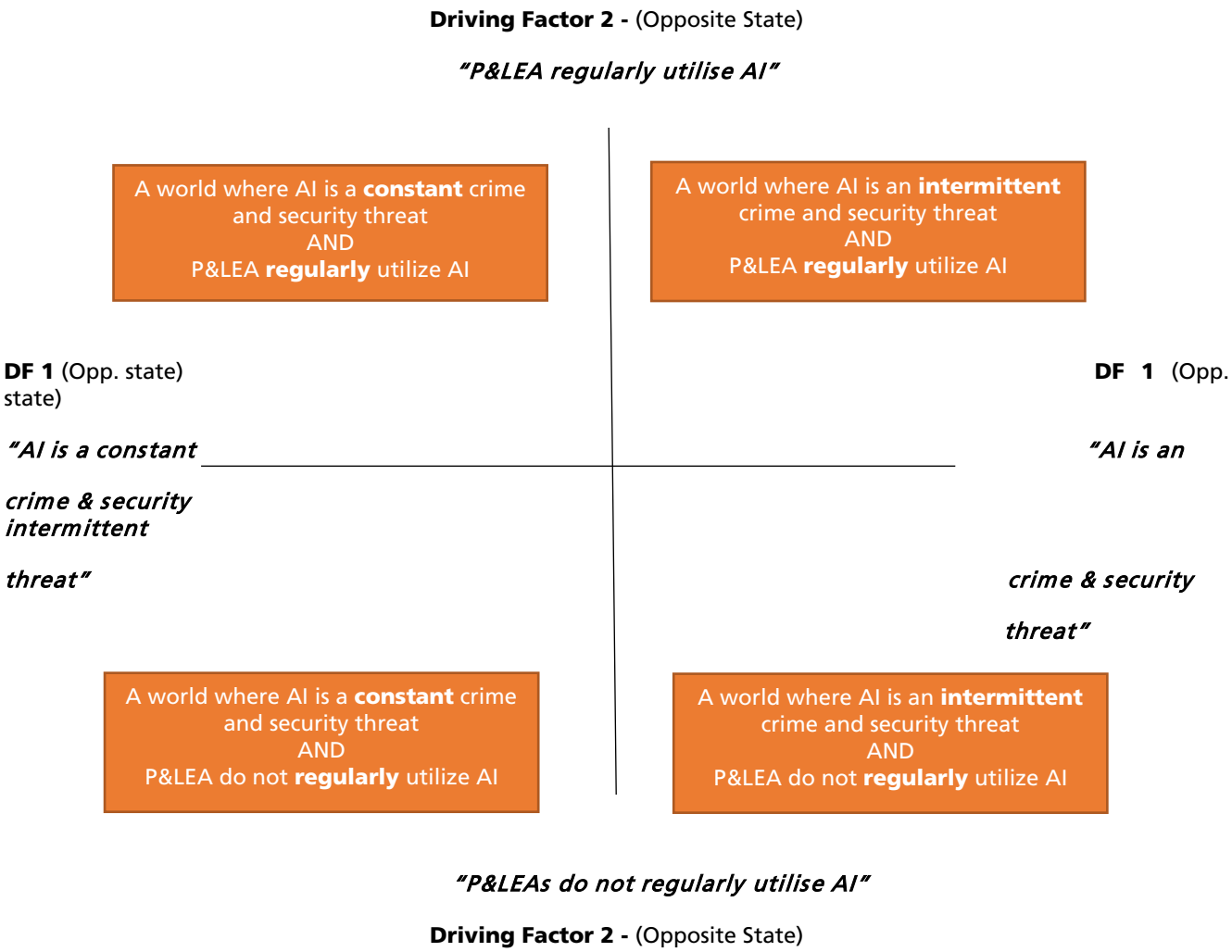
3.1.3 Creating 'four alternative scenario worlds' using a two axes method

When the two axes of the alternative future states, each with their polar opposite ends, were placed together as a cross with one perpendicular axis and one horizontal one, four alternative scenario worlds were generated (see Figure 6 below).

Figure 6: The four alternative scenario worlds created by the two axes of 'AI as a crime and security threat' and 'AI utilisation by Police and Law Enforcement Agencies (P&LEA)'
(Methodology after Noortman et al (2019), as adapted from Rockefeller Foundation (2010))

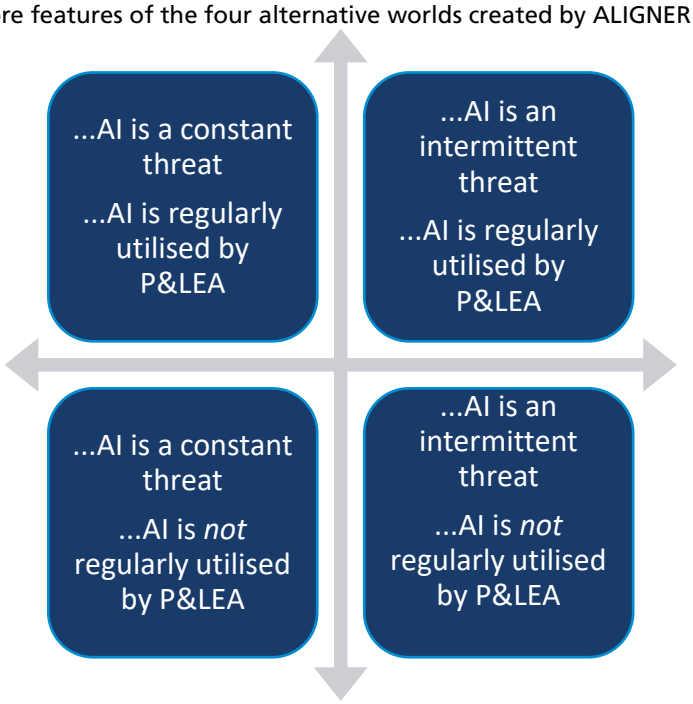
Driving Factor 1 (DF 1) - AI as a crime and security threat

Driving Factor 2 (DF 2) - Police and Law Enforcement Agency (P&LEA) utilisation of AI



The four alternative worlds created by using the two selected drivers of change are summarised below in Figure 7.

Figure 7: Summary of the core features of the four alternative worlds created by ALIGNER



Once the four worlds are created, they can be analysed to determine the core criteria and relevant factors that might be found within them.

3.1.4 Selecting the most appropriate alternative world for the ALIGNER archetypical scenario

Once the four alternative worlds defined by the two axes have been created, the methodology of Noortman et al. suggests that scenario narratives can be created to describe and investigate each of them.

However, the requirements of ALIGNER are better served by an alternative approach; out of the four alternative worlds, selecting a single world that is best suited to serve as the archetypical scenario. Consequently, the alternative world most applicable to act as the ALIGNER archetypical scenario is the world created from within the top left quadrant of the two axes grid. It best encapsulates the world created from the two drivers of change selected by ALIGNER that is, a world where 'AI is a constant threat' and 'AI is regularly utilised by P&LEA', as expressed below in Figure 8.

Figure 8: The ALIGNER Archetypical Scenario

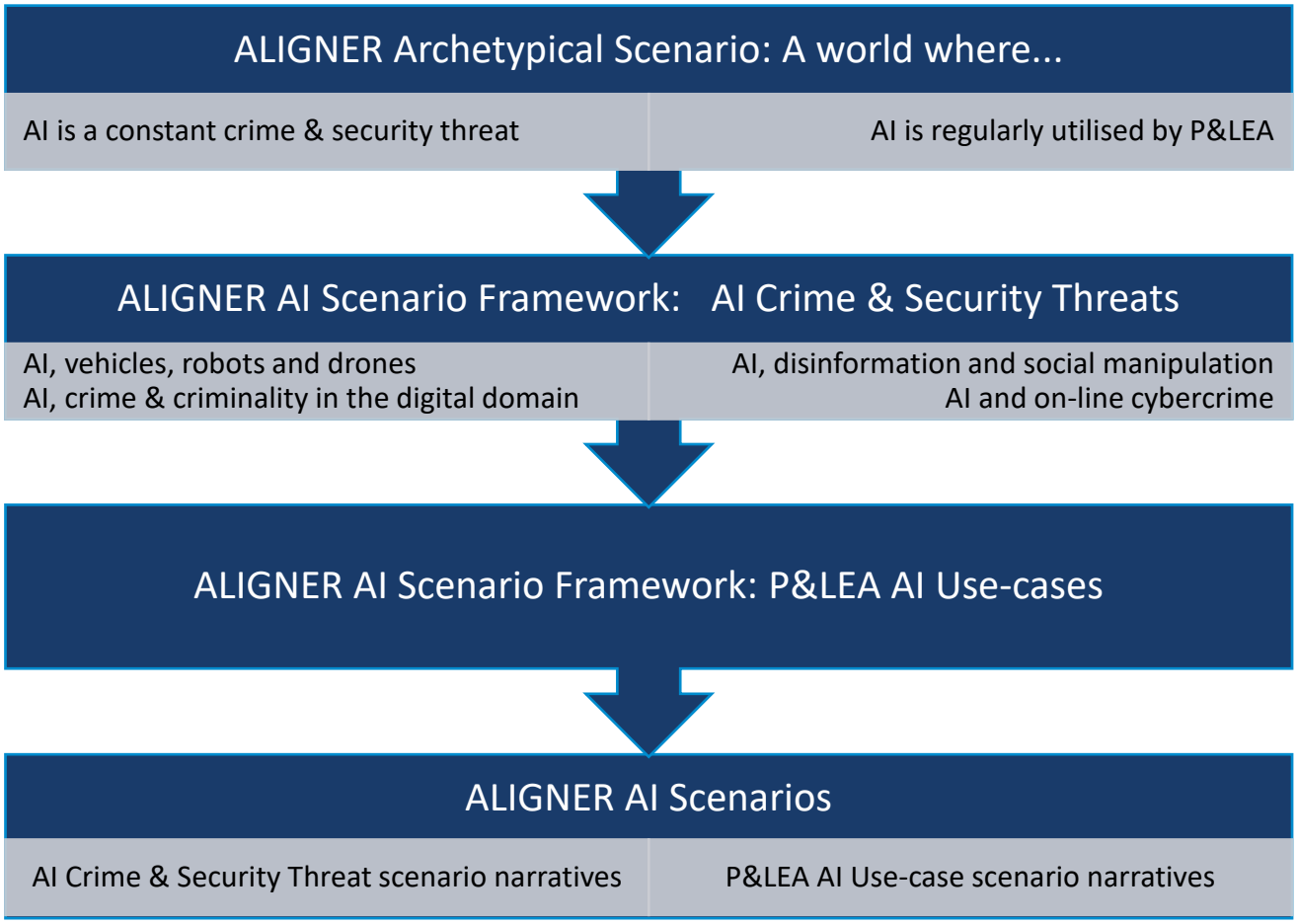
The ALIGNER Archetypical Scenario is a world where AI in one form or another is a constant crime and security threat to individuals, organisations and society, and where Police and Law Enforcement Agencies regularly utilise AI to carry out their responsibilities, duties.

From the perspective of ALIGNER scenario development, all four of the alternative worlds were considered but three of these were not considered suitable. For example, in one (bottom right) a judgement can be made that there is only a low probability of any near-future 'world' existing where AI is only ever an 'intermittent threat' and where AI is 'not regularly utilised' by P&LEA. Consequently, these three worlds will not be deliniated further.

4. The ALIGNER Scenario Framework: AI Threat scenarios and AI Use-case scenarios

The ALIGNER Archetypical Scenario characterises the chosen 'world' in which AI and P&LEA interact together and if required, it can be used directly to help generate specific scenario narratives. However, as part of the ALIGNER scenario structure, it can also be used in conjunction with the Scenario Framework and then, in turn, it can contribute to the creation of appropriate scenario narratives. This is shown in Figure 9 below.

Figure 9: The ALIGNER Scenario Structure



The first part of the Scenario Framework are the typology categories and sub-categories of AI Crime and Security Threats developed initially as a consequence of the first ALIGNER Workshop and encompassing all the examples raised and discussed at the workshop reflecting the perspective of 'AI as a Crime and Security Threat' (see Figure 10 below)

Figure 10: ALIGNER Scenario Framework for AI Crime and Security Threats: Typology Categories, sub-categories (with some examples)

Typology Category: ***AI, vehicles, robots and drones***

Weaponised or criminalised autonomous vehicles

AI controlled robots (Civil and militarised)

Weaponised or criminalised autonomous drones

Typology Category: ***AI, crime and criminality in the digital domain***

AI enabled fraud & forgery

AI data harvesting & exploitation

(e.g. as an enabler for acquisitive crime)

AI enabled social engineering

Typology Category: ***AI, disinformation and social manipulation***

AI enabled social engineering

(e.g. 'phishing', chatbots, hyper-personalised misinformation)

AI 'Deep Fakes' & Impersonation

AI data harvesting & exploitation

(e.g. as an enabler for crime within societal & political domains)

Typology Category: ***AI and on-line cybercrime***

AI disruption of AI systems

(e.g. for Denial of Information/Service, autonomous vehicles, etc)

Exploitation of AI capabilities

(e.g. learning and automation)

AI use in AI countermeasures

(e.g. evasion of detection)

AI data poisoning & data bias

(e.g. creation & exploitation of flawed AI)

The four highest-level, over-arching typology categories of AI Crime and Security Threats from Workshop 1 form part of the Scenario Framework and were also utilised in Workshop 2 to guide discussions relating to 'P&LEA capability enhancement needs' (see Figure 4 previously). However, the Scenario Framework also includes typology AI utilisation categories relating to the other dimension of AI; 'AI utilised in the service of P&LEA'.

These additional categories are:

Recognition and Identification of Individuals

Crime and threat; detection and prevention

Data and information handling processes

Digital forensics

Digital domain activity

Autonomous vehicles, robots and drones

The final element for inclusion in the Scenario Framework was the categories based on the core P&LEA functions of 'Prevention and Detection', 'Reaction and Response' and 'Investigation and Prosecution'. The category of 'Ancillary P&LEA functions' was devised to capture any other AI functions that may be required. After considering the discussions and outputs from both of the workshops, two groups of P&LEA AI utilisation categories most relevant to the Scenario Framework were selected and are summarised below in Figure 11.

Figure 11: ALIGNER Scenario Framework for AI P&LEA Use-cases: Categories

Typology Categories drawn from 'AI as a Crime and Security Threat'

AI, vehicles, robots and drones

AI, information and crime across the digital domain

AI, disinformation and social manipulation

AI and on-line cybercrime

Typology Categories drawn from 'AI in the service of P&LEA'

Recognition and Identification of Individuals

Crime and threat, detection and prevention

Data and information handling processes

Digital forensics

Digital domain activity

Autonomous vehicles, robots and drones

Prevention and detection capabilities

Reaction and response capabilities

Investigation and prosecution capabilities

Ancillary P&LEA capabilities

With the selection of the ALIGNER archetypical scenario and in the light of the typology categories applicable to each of the two types of scenario (AI Threat scenarios and AI Use-case scenarios), consideration can then be given to selecting specific scenarios and writing a scenario narrative for it.

5. Generating scenario narratives for AI Threat scenarios and AI Use-case scenarios

While the project requirement for ALIGNER is to create four AI 'crime scenarios', it is clear that concentrating only on the single 'crime and security threat' dimension of AI could lead to the neglect of the other, perhaps more important dimension of AI, namely where AI is utilised in the service of P&LEA. The approach taken by ALIGNER gives an opportunity to generate scenario narratives suitable for use in either 'AI Threat scenarios' or 'AI Use-case scenarios'.

Both types of scenario narrative ultimately originate from and are rooted in the Archetypical Scenario and Scenario Framework. The 'dual-type scenario' approach suggested can be used to envision and select from the multiple perspectives required to consider the practical aspects of AI in the context of policing and law enforcement.

For every scenario narrative required, one of these two types of AI scenario must be selected to act as a focal point. Once the scenario type is decided, the topic and title for the scenario and the key elements that are to be included in it can then be selected and finally, all of them can be integrated to produce the specific scenario narrative.

5.1 Narratives for AI Threat scenarios

An ALIGNER AI Threat scenario narrative will cover in broad outline the technological aspects of the selected AI, as described in the 'factsheets', impact and risk assessments that will form part of the output of ALIGNER WP3. It will also include how the AI is or could be utilised by a 'hostile actor' and hence will include the modus operandi (MO) plus the tactics, techniques and procedures (TTPs) involved.

If considered appropriate, the narrative could also include the technological aspects of any AI utilised by P&LEA in their reaction and response to both the 'hostile' AI and to the 'hostile actors' MO and TTPs. Also, within the overall scenario, the ethical, legal and social implications of any P&LEA response where AI is utilised could be identified and explored.

5.2 Narratives for P&LEA AI Use-case scenarios

The development by ALIGNER of relevant P&LEA AI Use-case scenarios and their accompanying narratives will make as beneficial a contribution to the overall aim of ALIGNER as its AI Threat scenarios will. Their main focus is likely to be on the P&LEA and its involvement with the AI technology as a capability it either controls or is able to utilise. While there are similarities, the are also differences in what could be included in P&LEA Use-case scenario narratives.

As in 5.1 above, they will include the technical aspects of the selected P&LEA AI technology and also, the P&LEA utilisation of the AI technology. This includes the P&LEA MO (a broad outline only) and the issues surrounding the deployment and employment of the AI.

The ethical, legal and social implications of the P&LEA utilisation of the AI technology or its acquisition could be identified and explored. This aspect is even more important in P&LEA AI Use-case scenarios as all the decision making, technological control and overall accountability for the AI rests with the P&LEA implementing the AI response or capability.

A Use-case scenario can include an AI threat-based/hostile element if appropriate and if it does so, it must also address the technological aspects of the P&LEA AI under examination and the MO and TTPs of the 'hostile actor'.

Finally, a variation of a Use-case scenario narrative could be used to explore the P&LEA AI requirements capture and procurement processes, with their attendant ethical, legal and social implications.

Certain potential topics for scenarios arising from both AI as a crime and security threat or AI in the service of P&LEA at first may appear to be identical or at least very similar e.g. the AI typology category of 'AI, vehicles, drones and robots'. Despite this, the scenario narrative written for an AI threat scenario will differ in significant ways from a Use-case scenario narrative, even if the AI technology in both the scenarios is the same or similar.

The most noticeable differences will arise from whether the AI technology is utilised by a 'hostile actor' or by a P&LEA. In each case, their motive, intent and overall aim will be significantly different and in turn, this will impact on their MO and TTPs. Another clear difference will be the elements of the scenario narrative relating to the ethical, legal and social implications. For example, the decision making, technological control and overall accountability for the AI and its use rests with the P&LEA employing it. Hostile actors using AI technology have no such constraints.

5.3 The first ALIGNER scenario

The narrative for the first ALIGNER Scenario will be published in September 2022 as part of the first iteration of deliverable D5.3, *"Research Roadmap for AI in support of Law Enforcement and Policing"*. It will be an 'AI Threat scenario', drawn from the AI Crime and Security Threat category of 'AI, disinformation and social manipulation' and will also include aspects from the category of 'AI in the service of Police and Law Enforcement'. Work on the scenario narrative is currently underway.

As called for in the Description of Work, the starting point for all four Scenarios will be developed within WP3, task 3.3 'AI Technology Risk Assessment'. WP2 will write the scenario narratives and consult with the Advisory Boards before each one is published by WP5 in the appropriate 'research roadmap' deliverable as part of task 5.3.

6. Conclusion

Prior to the first Workshop, preliminary desk research showed that the domain of AI and its impact on Policing and Law Enforcement was not only vast but amorphous as well. Currently, there appear to be few well-defined boundaries or certainties within it, yet it is clear that the growing influence of AI has already begun to have an impact on society. Policing and law enforcement is a fundamental element of any democratic society based upon the 'Rule of Law' and consequently, the advance of AI will have profound implications on the ways in which police and law enforcement agencies need to adapt to these changes. They must do this now, and continue to do so into the future as AI and the ways it is used change and evolve.

Beginning with the discussions during ALIGNER Workshop 1 and reinforced during those of Workshop 2, a sense emerged that there is a gap in many Police and Law Enforcement Agencies, not just between their current technological capabilities and their future AI capability needs but more worryingly, between their current technological capabilities and whether AI needs to be considered today as a necessary requirement or deferred into the future.

There was a feeling that, generally (and with notable exceptions), P&LEA leadership corporately may not yet be convinced of the fundamental changes and impact AI can and will bring to their organisations and the duties they perform and further, that addressing the issue is much wider and more complex than just determining what kinds AI technology are desirable, available and affordable. On the other hand, there are an increasing number of cases where the 'rush for AI' by P&LEAs has resulted in ill-thought through AI technology purchases that have led to its problematic, and even unlawful, employment. The need for well-founded and impartial practical guidance and advice for P&LEAs is already clear.

Consequently, without it, many P&LEAs have not engaged fully with the basic technological, human, legal and ethical factors that must be addressed in order to best identify, procure and utilise AI, or how AI technology can also pose crime and security threats that their organisations have a duty to respond to. In the United Kingdom, the Chair of the Strategic Review of Policing in England and Wales (due to publish its report shortly) commented in a recent speech that *"In the digital age, where the dark web is often the new crime frontline, it can feel like a contest between a Betamax police force and blockchain enabled criminals."*[7] How P&LEAs as organisations can effectively bridge this technological gap is still not clear but there is no doubt that there will be calls for them to utilise AI technology in order to do so.

This deliverable has set out how ALIGNER has devised a concept, structure and process to develop AI-centric scenarios as a means of contributing to the overall project aim. All ALIGNER AI scenarios and scenario narratives will reflect the dual nature of AI in the context of policing and law enforcement, where AI technology can be a crime and security

[7] 'Police review chief says 'Betamax police' stuck in the past', Rajeev Sayal, *The Guardian*, February 21st, 2022 https://www.theguardian.com/uk-news/2022/feb/21/police-review-chief-says-betamax-police-stuck-in-the-past Accessed: 24.2.22

threat and simultaneously, it can be utilised by Police and Law Enforcement Agencies as a means to increase their capabilities and drive improvements in their effectiveness.

Finally, it is important that the ALIGNER scenarios are produced to meet two criteria; they are credible, relevant and rooted in the current realities facing Police and Law Enforcement Agencies and at the same time, they enable the thinking required to encompass emerging and foreseeable developments in AI, its potential impacts on policing and law enforcement capabilities and functions, and more widely, upon society itself.

References

'The Malicious use of Artificial Intelligence: Forecasting, prevention, and mitigation' Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., Dafoe, A., Scharre, P., Zeitzoff, T., Filar, B., Anderson, H., Roff, H., Allen, G. C., Stein- hardt, J., Flynn, C., hÉigeartaigh, S. Ó., Beard, S., Belfield, H., Farquhar, S., Lyle, C., Crootof, R., Evans, O., Page, M., Bryson, J., Yampolskiy, R., & Amodei, D. *arXiv.org*, vol. cs.AI. (2018) https://arxiv.org/pdf/1802.07228.pdf

'AI-enabled future crime'
M.Caldwell, J.T.Andrews, T.Tanay and L.D.Griffin
Crime Science (2020)
https://link.springer.com/content/pdf/10.1186/s40163-020-00123-8.pdf

'Imaginative Scenario Planning for Security and Law Enforcement Organisations' Noortmann, M., Koning, J., Vervoort, J. and Hoofd, I.; (2019)
https://crestresearch.ac.uk/resources/imaginative-scenario-planning-toolkit/

'Police review chief says 'Betamax police' stuck in the past', Rajeev Sayal, *The Guardian*, February 21st, 2022
https://www.theguardian.com/uk-news/2022/feb/21/police-review-chief-says-betamax-police-stuck-in-the-past

'Grounded Theory Methodology: An Overview' Strauss, A. & Corbin, J. (1994). In N. Denzin & Y. Lincoln *Handbook of Qualitative Research. 1st ed.* (pp 273-284) and Warnes, R. (2009) 'Grounded Theory'. In Ling, T. & Villalba van Dijk, L. *Performance Audit Handbook: Routes to Effective Evaluation.* Santa Monica, Calif., Rand Corporation (Chpt. 10)
https://www.rand.org/pubs/technical_reports/TR788.html

Annex A

Data collection grid showing ALIGNER typology categories & sub-categories for examples of 'AI as a Crime and Security Threat'

AI Crime & Security Threat Categories	AI & Vehicles, Drones & Robots			AI Crime & Criminal Activity in the Digital Domain			AI, Disinformation & Social Manipulation		AI & On-line Cybercrime		
AI Crime & Threat Sub-Categories	Weaponised/ criminalised autonomous vehicles	AI controlled robots (Civil & militarised)	Weaponised/ criminalised autonomous drones	AI enabled fraud & forgery	Exploitation of AI capabilities	AI enabled social engineering	AI 'Deep Fakes' & Impersonation	AI data harvesting & exploitation	AI disruption of AI systems	AI use in AI countermeasures	AI data poisoning & data bias
AI threat examples											
Terrorist repurposing commercial AI systems	X	X	X								
Example 2											
Example 3											
Example 4											
Example 5											
(Continued)											

Annex B

Data collection grid showing ALIGNER typology categories for examples of 'AI utilised in the service of Police and Law Enforcement Agencies'

Police & LEA AI Capability Categories	Recognition & Identification of Individuals	Crime & threat, detection & prevention	Data & Information handling processes	Digital forensics	Digital domain activity	Autonomous vehicles, robots & drones	Prevention and detection capabilities	Reaction & response capabilities	Investigation & prosecution capabilities	Ancillary Police & LEA capabilities
P&LEA AI utilisation examples										
Facial recognition	X	X		X	X		X	X		
Example 2										
Example 3										
Example 4										
Example 5										
Example 6										
(Continued)										