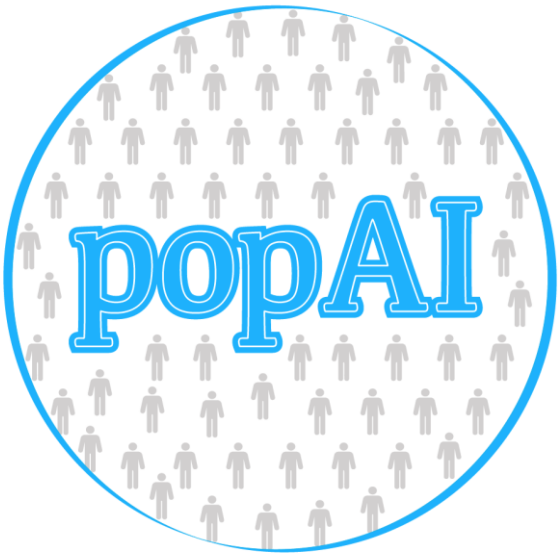


A European Positive Sum Approach towards AI tools in support of Law Enforcement and safeguarding privacy and fundamental rights

D3.2: Report on Citizen discourses and attitudes towards controversies

Grant Agreement ID	101022001	**Acronym**	popAI
Project Title	A European Positive Sum Approach towards AI tools in support of Law Enforcement and safeguarding privacy and fundamental rights		
Start Date	01/10/2021	**Duration**	24 Months
Project URL	https://www.pop-ai.eu/		
Submission date	11/10/2022		
Nature	R = Document, report	**Dissemination Level**	PU = Public
Authors	Pinelopi Troullinou, Ilaria Bonavita, Fabienne Ufert (TRI) Francesca Trevisan (ERI) Simeon Stoyanov (ECAS)		
Contributors	Claire Morot – Sir (ECAS) Anastasios Drosou (CERTH)		
Reviewers	(TRI), (NCSRD)		

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 101022001.

Executive Summary

Technology has always been used to support crime investigation and even crime prediction. It is of no surprise that in the digital era, the employment of AI in the civil security domain is ever-increasing and broadly accepted. However, the surveillance nature and the general mistrust to AI as well as the lack of transparency amplify concerns regarding the potential impact and harm to the society and fundamental rights. It is very important to understand the public discourses around the employment of AI in the sensitive area of law enforcement to increase awareness and consequently boost trust. Equally, the identification of key influencers can assist the dissemination of recommendations for ethical use of AI supporting popAI's vision on a unified European view based on common values and principles. This deliverable contributes to this objective drawing upon empirical research namely media analyses on data from the popular social media platform Twitter.

The report consists of four main sections. Section 1 introduces the deliverable and presents its objectives and approach.

Section 2 provides a more theoretical background discussing both the historic and emerging discourses on the advantages and concerns/risks related to AI employment for law enforcement and judicial purposes. Particular attention is being paid to relevant discourses by EU institutions and bodies, civil society organisations and the lay public. The need for transparency and accountability to minimise serious risks on fundamental rights is a common discourse.

Section 3 describes the three types of media analysis employed, namely social media listening, network analysis and discourse overview via natural language processing and presents the main findings. Social listening has been employed to understand 'at a glance' the key topics discussed online as well as the sentiment with which these areas are discussed, positive, negative, neutral. Network analysis and discourse overview via NLP focused on Twitter examining the key influencers in the public discourses around the use of AI and how diverse stakeholders discuss AI controversies.

Section 4 discusses the findings of the media analyses confirming the importance of the stakeholders identified in the controversy ecosystem mapping (Task 3.1). The findings also showed the under-connectedness of civilian and activist group accounts with the most influential accounts belonging to established organisations and political figures, such as the European Commission, the European Parliament or the United Nations.

Section 5 provides the main conclusions emerging from the theoretical background and the empirical analysis presented in the report. The need for transparency and accountability is discussed and the dissemination of a unified European view to promote trust in AI use for Law Enforcement.

Table of Contents

List of Figures

List of Tables

List of Terms & Abbreviations

Abbreviation	Definition
TRI	Trilateral Research Limited
ERI	ETICAS Research and Innovation
CERTH	Centre for Research and Technology Hellas
ECAS	European Citizen Action Service
NCSRD	National Centre for Scientific Research “Demokritos”
AI	Artificial Intelligence
NLP	Natural Language Processing
LEA	Law Enforcement Authorities
EU	European Union
API	Application Programming Interface
GDPR	General Data Protection Regulation
CSA	Coordination and Support Action
CSO	Civil society organisation
EP	European Parliament
MEP	Member of European Parliament
AFSJ	Area of freedom, security, and justice
AR	Augmented Reality
EC	European Commission
HLEG	High Level Expert Group on AI
ETIAS	European Travel Information and Authorisation System
WEF	World Economic Forum

1 Introduction

popAI is a 24-month Coordination and Support Action (CSA) project funded by Horizon 2020 undertaken by a consortium of 13 partners from 8 European countries. The core vision of the project is to boost trust in AI by increasing awareness and current social engagement, consolidating distinct spheres of knowledge, and delivering a unified European view and recommendations. In this context, it is key to explore the perspectives of different stakeholders involved in the design, development, and deployment of AI in the security domain.

1.1 Scope and objectives of the deliverable

This deliverable (D3.2) entitled *Report on citizen discourses and attitudes towards controversies* is mainly of empirical nature and draws upon three types of media analyses, namely social media listening, network analysis, and discourse overview via natural language processing. The objective is to shed light on the way the public discusses the use of Artificial Intelligence (AI) in support of law enforcement and to identify key influencers that communicate these topics. In the context of our work, influencers are understood as social media accounts that are well-connected in a network of interactions having many two-way connections to other accounts. This means that they act essentially as a kind of nexus, or central hub, through which information can flow.

Three media analyses approaches were employed namely, social (media) listening analysis, network analysis, and discourse overview via NLP. Social media listening analysed data from the open web repository, CommonCrawl, for the period 2013-2021. The analysis indicated which areas of AI usage in LEAs generate most discourses on the internet using the main areas that emerged from Task 3.1 and developed as a framework for the crowdsourcing platform (Task 3.3); biometric identifiers, predictive policing, police hacking, cyber operations, decision making in the justice systems. The analysis also allowed the understanding of the sentiment with which these areas are discussed; positive, negative, neutral. This deliverable discussed the methodology the findings and analysis will be provided in the deliverable entitled *Stakeholder attitudes, priorities, and recommendations for addressing AI in the security domain* (D3.3, due in March 2023).

Network analysis and the discourse overview via NLP collected data from Twitter, using keywords that emerged mainly from the mapping of the controversy ecosystem of AI tools in the security domain (T3.1) [1]. The network analysis on Twitter data indicated the way networks that discuss AI use in LEAs are formed and provided the key influencers in these networks. The discourse overview via NLP identified the most commonly occurring words and words pairing indicating the discourses around AI.

It is important to understand how and whether human factors and ethical, societal, and legal aspects of using AI in security domain are discussed *in* and *by* the public to address them respectively and increase trust. In addition, identifying what kinds of stakeholders are key nodes and influencers in public debates allows the project partners to invite them in future activities.

[1] D3.1 "Map of AI in policing innovation ecosystem and stakeholders" submitted on 31st March 2022

1.2 Approach for Work Package and Relation to other Work Packages and Deliverables

The deliverable is the outcome of the Task 3.2 *Understanding citizen discourses around AI and security controversies* building upon the controversy and ecosystem mapping reported in D3.1 *Map of AI in policing innovation ecosystem and stakeholders*. The findings allow the project to identify a 'nudge' strategy for proactively addressing public discourse that will support project's future activities, and more specifically:

- Task 3.4 Engaging LEAs and relevant experts through policy labs
- Task 3.3 Crowdsourcing stakeholder attitudes and pro-active solution ideations
- Task 5.1 Community building and ecosystem engagement activities

Furthermore, the understanding of the public discourses around the use of AI in the security domain will feed into WP4 *The pandect of recommendations for the ethical use of AI for LEAs*.

1.3 Structure of the deliverable

The remainder of this deliverable is organised as follows:

Section 2 presents an overview regarding the public discourses around the use of AI in the security domain.

Section 3 describes the media analyses approaches employed, namely social listening, network analysis, and discourse overview via NLP and illustrates the findings.

Section 4 discusses the findings of the media analyses.

Section 5 draws a conclusion and provides insights on how the findings can be useful in the overall objective of the popAI project.

2 Public discourses on AI in security domain

This section discusses the public discourses on AI in the context of civil security. Firstly, it provides an overview of the history and current state of the art of the AI use in the security domain. It also illustrates the emerging discourses on the advantages of AI employed for law enforcement and judicial purposes as well as emerging risks. Following, it discusses the main public discourses around the topic. For example, much of the public discussions on the use of AI in the security domain has depicted AI as a new form of discrimination and intrusion on citizens' fundamental rights. This report focuses on the understanding of citizens' discourses on AI use for civil security purposes. Therefore, attention is paid to EU institutions and bodies, civil society organisations, and lay public.

2.1 AI use in the civil security domain

The use of technology in the security domain is not novel. Technology has always been used to support crime investigation and even crime prediction as it was briefly discussed in the popAI's report "Map of AI in policing innovation ecosystem and stakeholders" (D3.1). In this section, an EU-focused historical perspective on the use of new technologies by LEAs will be offered, illustrating the emerging potentials, as well as overall risks.

In this report, we follow the broad definition of AI as adopted in previous reports of popAI[2]; including **any technology that is based on the digital processing of information to produce an outcome.** Legal and policy debates on AI in the EU stem from the EU Digital Single Market agenda and emerged in connection to 'big data' and regulating the processing of both personal and non-personal data (Gonzalez Fuster, 2020)[3]. Today, the idea that AI can also be of substantial use in the field of law enforcement and criminal justice is broadly accepted and endorsed at the EU level (EC, 2019), and the trend towards using automated processing techniques and algorithms in crime prevention and the criminal justice system has been described as generally growing for already a number of years (Gonzalez Fuster, 2020). But already in 2014, the European Commission noted that, although digitisation of public services opened up new opportunities to optimise data analysis, the "reported use of similar technologies for surveillance purposes, by public or private actors, is liable to feed concern and reduce trust in the digital economy among individuals and organisations", deserving thus special attention (EC, 2014). Debates on the benefits and risks of developing and deploying new technologies in law enforcement are omnipresent but often suffer from persistent ambivalences on the role of 'ethics' in relation to the safeguarding of fundamental rights, generating not only lack of conceptual precision but uncertainty as to whether effective protection of fundamental rights will be delivered (Gonzalez Fuster, 2020). The field of law enforcement and criminal justice is particularly sensitive, as it touches upon core issues of the relation between the individual and the State. There is a broad consensus on the fact that reliance on AI systems in this area can substantially impact EU fundamental rights and values, such as democracy itself (Gonzalez Fuster, 2020)[4].

AI functionalities and applications in the security domain are broad (see D2.1) but the themes that have emerged as the major ones in the EU are predictive policing, facial recognition, AI and criminal

[2] D3.1 discussed the definition of AI in civil security (Section 2.1).

[3] The relationship between AI and legal frameworks around data is further discusses in D2.2 *report Legal casework taxonomy: emerging trends and scenarios*, submitted in September 2022.

[4] D2.2 section on fundamental rights lists the fundamental rights that can be impacted by AI

justice, and AI and borders (Gonzalez Fuster, 2020)[5]. As for predictive policing, the origins might be traced back to experiments of computer-assisted crime control in the 1970s, and although prediction in the area of crime and punishment has been developed and discussed during many more decades (Harcourt, 2008), the term became eventually connected to the rise of 'big data' (Wilson, 2017). Prediction as an extensive trend has noticeably marked security developments globally and in Europe over the past decade at least, typically associated with prevention. Predictive policing initiatives have emerged in different Member States, mainly since the beginning of the 2010s, e.g., initiatives involving automatic license plate control systems (Gonzalez Fuster, 2020). The expansion of police use of algorithms in the recent years has been connected to three main factors: austerity measures that limit resources and push towards apparently more economically efficient new solutions; an increasing perception that police ought to adapt a preventative posture, placing emphasis on anticipating harm; and an increase in the volume of complexity of data available, necessitating increasingly sophisticated processing tools (Gonzalez Fuster, 2020).

The deployment of facial recognition in public spaces, including in test mode, has generally provoked a variety of reactions in the EU. For example, in 2017, the public was said to be divided when a field test of facial recognition programme was launched at Berlin's railway stations (Fuerstenau, 2017). Since then, the use of facial recognition has been declared invalid much more often than it has been allowed in many EU countries (Gonzalez Fuster, 2020).[6]

The origins of AI in criminal justice domain may be traced back to the emergence of 'computerised justice' as a trend aiming at improving the accountability and predictability of judicial decisions, by reducing discretionary powers in the hands of judges regarded by some, in some circumstances, as excessive (Franko Aas, 2005). However, the use of predictive tools by judges in criminal trials in Europe has been described as very rare, at least until recently (Ronsin & Lampos, 2018).

As for the use of AI in border management, security and border management have been major drivers for the growth of both centralised and decentralised information systems in the area of freedom, security and justice (AFSJ) for many years already (EC, 2016). The interest among public authorities in data to be processed for migration and asylum purposes appears to be increasing in many Member States and now, large scale information systems in the AFSJ increasingly incorporate algorithmic decision-making (Gonzalez Fuster, 2020).

Overall, the use of surveillance technologies – and most AI systems in the security domain are surveillance-based – by LEAs is increasingly met with societal mistrust and scepticism due to the lack of transparency as to the amounts of data that are used and the implications for the fundamental rights of citizens. Additionally, the use of AI in law enforcement bears risks to privacy and data protection but also other challenges, notably related to non-discrimination. D3.1 discussed concerns and specific campaigns against AI technologies such as facial recognition used by LEAs and are of high risk to fundamental rights. Civil organizations have also expressed concerns regarding ambiguous research projects that might raise potential risks. The need to strike a balance between developing

[5] D3.1 followed a more detailed taxonomy of the civil security domain structure to investigate controversial cases of AI use namely, Crime prevention, Crime Investigation, Migration, Asylum, and Border control, Administration of Justice, Cyber Operations for law enforcement, and Training.

[6] Other controversial cases regarding the use of facial recognition for LEA purposes have been discussed in D3.1

cutting-edge, AI-based technologies and regulating data-driven AI technologies, including robust safeguards for citizens, is ever more important. As an example, the EU-funded Horizon 2020 project DARLENE investigates and develops innovative augmented reality (AR) tools to improve the situational awareness and capacity of law enforcement agencies (LEAs) to make well-informed and rapid decisions in real-time when responding to live criminal and terrorist incidents. DARLENE partners apply an iterative and inclusive ethics-by-design approach to proactively introduce ethical principles in the development of the DARLENE technology and other tasks.

2.2 Public discourses on the use of AI in security

The use of AI in the civil security domain as discussed above is employed in the name of security and efficiency. However, the use of AI systems by LEAs raises great public debates mainly on the risks it bears. Nissenbaum (2009), a prominent privacy scholar, has argued that socio-technical systems, like AI-based technology in law enforcement, have often been introduced "without a careful evaluation of harms and benefits, perturbations in social and cultural values". It is often the case that no effective anticipatory evaluation and, thus, control is exercised over the ethical and societal implication of these novelties. What is clear is that, when it comes to technology with surveillance capabilities for example, "the risks […] go well beyond anything that quests for 'privacy' or 'data protection' can cope with on their own" (Lyon, 2003).

AI systems in the security domain are Surveillance-Orientated Security Technologies (SOSTs) meaning, "technologies which collect information about the general population to monitor the activities of potential suspects and to prevent criminal acts from occurring" (Esposti & Gomez, 2015, p. 437). Surveillance is an ambiguous practice, creating both positive and negative effects (Lyon, 2003). Innovations in technology have facilitated and improved the possibilities for humans to carry out surveillance activities, making them more efficient and effective, in response to the growing complexity of contemporary crime and security threats. At the same time, questions arise as to how ethically and practically acceptable the use of these technologies is given the new opportunities for potential interference with human rights and ethical values brought by such tools (Esposti & Gomez, 2015). Failure to implement effective surveillance systems can hamper an effective response to public security threats. However, heightened levels of surveillance may lead to undesirable effects, which in turn have often triggered public resistance towards and backlash against surveillance technology (Cohen, 2017).

When referring to the ethics of surveillance, it is worth considering the research performed in the FP7 SURVEILLE project (SURVEILLE, 2020). Based on scenarios developed to represent real or possible usage situations, the SURVEILLE project considered and addressed issues related to surveillance technologies from a range of angles, generating valuable input for surveillance studies. Considering the findings under that project, it can be said that there are three distinct types of effects of surveillance technology seen from an ethical perspective, namely intrusion, error and discrimination, and trust and chill. Information on these types of effects is provided below:

1. Intrusion, consisting of the penetration of a person's "private interests" (e.g. relaxation, intimacy, family life, association and independent thought);
2. Error and discrimination, which could lead to certain wrongs, such as false arrests in a surveillance context, ultimately leading to other harms, such as social exclusion or disadvantage;

3. Damage to trust, either in the authorities, which could ultimately be bad for democracy, security, and chill, or between citizens.

Against this background, the following sub-sections summarise the public discourse around AI use by LEAs at different levels.

EU institutions and bodies

European governmental bodies (e.g., European Commission (EC), European Parliament (EP)) aim to develop trust around the use of AI (EC, 2019) and acknowledge some potential benefits of employing AI for security purposes, such as the ability to detect, prevent and reduce criminal activities. To that end, the EC has been very active, e.g., through the appointment of its "High Level Expert Group on AI" (HLEG). The HLEG has produced its landmark "Ethics Guidelines for Trustworthy AI" (HLEG, 2019). These Guidelines form part of a broader vision to embrace a human-centric approach to AI, intended to make Europe a global leading innovator in ethical, secure, and cutting-edge AI. Accordingly, the guidance document strives to facilitate and enable "Trustworthy AI made in Europe" that will enhance the wellbeing of European citizens (HLEG, 2019). Furthermore, the European Group on Ethics in Science and New Technologies, an independent and multi-disciplinary advisory body to the EC with regard to ethical, societal and fundamental rights issues related to emerging technologies (EC, 2022), has produced an opinion on the ethics of security and surveillance technologies (EC, 2014).

However, EU institutions and bodies also highlight the risks that AI applications entail for the protection of fundamental rights and democracies (e.g., Section 3.5 AI Act), the challenge of defining legal responsibilities and liability for potential harm and the danger of employing AI on promises that do not always hold true. In 2020, the EP criticised that deliberations on a possible need for an AI regulatory framework have primarily surfaced and been developed in a Digital Single Market context. However, these debates have only touched upon the challenges of AI in law enforcement and criminal justice in a limited manner (Gonzalez Fuster, 2020). The EP's Draft Report on AI in criminal law and its use by the police and judicial authorities in criminal matters (LIBE, 2020/2016(INI)) includes a call for a moratorium on the deployment of facial recognition systems for law enforcement, as well as a set of recommendations, including the establishment of compulsory fundamental rights impact assessments and the creation of a clear and fair legal regime for assigning legal responsibility for the potential adverse consequences produced by advanced digital technologies in the field (Gonzalez Fuster, 2020).

Moreover, in an EP Resolution adopted in October 2021 (EP, 2021), Members of Parliament (MEPs) asked for strong safeguards when AI tools are used for law enforcement, to prevent discrimination and the abuse of fundamental rights. The Resolution emphasised how opaqueness in AI-based decision making, and the errors and discrimination inherent to algorithms pose serious risks for citizens' privacy and protection of personal data, freedom of expression, presumption of innocence, right to an effective remedy and a fair trial. In this context, the EP called for transparency, explainability and traceability of AI systems and AI-powered decision making, as well as periodic audits and compulsory fundamental rights assessment to be conducted before the use of any AI system for law enforcement and judiciary purposes.

Consequently, societal and ethics impact assessments have become a common prerequisite for security research under the Horizon 2020 and Horizon Europe frameworks. For example, the Horizon 2020 project DARLENE recognises the need for, and conducts an assessment of the societal and

ethical impact of research outcomes and provides measures, tools and procedures such that end-users, citizens and society as a whole can accept and use cutting-edge, AI-powered Augmented Reality (AR) solutions in the fight against crime and terrorism

Civil society

Civil society organisations (CSOs) play a key role in shaping the discourse on AI in the security domain, with a focus on what needs to be done to protect citizens' fundamental rights and civil liberties. Civil society organisations emphasise how the lack of transparency and safeguards put in place for the deployment of AI in policing and criminal justice results in practices that discriminate and endanger fundamental rights. For example, CSOs emphasise how predictive policing reproduces and reinforces discrimination by using datasets that do not reflect actual criminal activity but policing priority surveillance (EDRi, et al., 2021). They highlight how biometric recognition puts people's privacy and freedom at stake and can lead to mass surveillance (EDRi, et al., 2021), how emotion recognition technologies are based on scientifically biased theories and affect human rights (EDRi, et al., 2021), and how AI powered technologies used for border control and migration management lead to the severe intrusion on fundamental rights and take a toll on the most vulnerable.

Documents that reflect the development in relation to predictive policing, facial recognition, AI and criminal justice, and AI and borders (including a reflection on the European Travel Information and Authorisation System, ETIAS) have already generated significant controversies and have led civil societies to call for a better prevention or mitigation of associated risks, both in the EU and beyond (Gonzalez Fuster, 2020).

Citizens

Evidence of public understanding and discourse on AI use in the security domain is scant and indicates citizens' preoccupations for how AI will impact their rights and for the low level of transparency and explainability of AI applications. A survey conducted by Ipsos for the World Economic Forum revealed that globally only 37% of respondents thought that AI would improve their freedoms and legal rights (WEF, 2022). Furthermore, the survey showed that trust in AI was higher in emerging countries (e.g., Saudi Arabia, India, China), as opposed to high-income European countries (e.g., Italy, Germany, France, Belgium, Sweden, the Netherlands). A qualitative study by Drobotowicz et al. (2021) investigating the use of AI in the public service suggested that citizens want a transparent and explainable public use of AI with control for personal data usage and the involvement of humans during the deployment of AI applications. Recently, the Accountability Principles for AI (AP4AI) research group conducted a citizen consultation with respondents from 27 EU member states, Australia, UK, and USA, to develop an AI accountability framework for policing, security and justice (Akhgar, et al., 2022). Overall, participants agreed or strongly agreed that AI use by police can benefit the society (72.1%). When looking at specific AI applications, 89.7% of respondents agreed or strongly agreed that AI should be used for the protection of children and vulnerable groups, and 88.6% agreed that AI should be used to predict crimes before they happen. Yet, participants expressed important privacy and bias concerns. Half of the sample was concerned of how AI would make their online and offline information more open to police scrutiny and about one in three participants (36.2%) worried about the negative effect of biased decisions. At the same time, the overwhelming majority of participants expected that police should be held accountable for the way they use AI (92%) and for its consequences (92%). In the qualitative answers collected by the survey, citizens expressed the importance of establishing binding laws to protect privacy and asked for more transparency on 1) how AI is used and its effect, and 2) police operations that involve AI and asked for evidence that show how AI for policing translates in positive change.

The employment of AI in the sensitive fields of law enforcement and criminal justice promises greater security and efficiency while generating also great risks on fundamental rights, as well as democracy itself. The section having provided an overview of the use of AI in security domain, it discussed the relevant public discourses paying attention to the EU institutions and bodies, civil society organisations, and lay public. Studies showed that EU citizens, especially from high-income European countries have low levels of trust in AI. This might be partly the reason why transparency has been a key principle in public discourses. The civil society organisations and the lay public emphasised on the importance of transparency regarding the AI use in policing so to address discrimination and abuse of fundamental rights. Initiatives of European institutions such as the appointment of EC "High Level Expert Group on AI" (HLEG) and the European Group on Ethics in Science and New Technologies, an independent and multi-disciplinary advisory body are very active on promoting measures to ensure ethical use of AI. Members of the European Parliament also report risks emerging from the deployment of technologies that use biometrics, such as facial recognition systems and take actions to address these issues. The next section analyses the networks of the public discourses around the use of AI and the key influencers shedding also light to the most discussed topics around the use of AI in civil security drawing upon empirical research.

3 Media analyses on citizen discourses around AI and security controversies

Media analyses methods were deployed to explore discourses around AI use in security domain and identify the respective influencers in the public sphere, namely social (media) listening, network analysis and discourse overview via NLP. In particular, social listening has been employed to understand 'at a glance' the key topics discussed online as well as the sentiment with which these areas are discussed, positive, negative, neutral. Network analysis and discourse overview via NLP focused on Twitter examining the key influencers in the public discourses around the use of AI and how diverse stakeholders discuss AI controversies.

3.1 Methods and objectives

Social (media) listening is a term most frequently used in marketing to denote a process through which one can identify what is being said about a certain brand, product, service, or topic. Agencies and companies used it to acquire competitor intelligence, see how the public perceives their new product and follow the latest industry trends. When it comes to social listening in social sciences sphere, the instrument serves a similar purpose - to let researchers understand how societal trends about a certain topic progress over time. The main objective of the social (media) listening exercise in the context of Task 3.2 is to observe the content as well as he tone (positive, negative, neutral) as of the online discourses to understand how the use of AI by LEAs is perceived by the public.

Network analysis and discourse overview via NLP serve to assess the way information dissemination networks are formed and to develop a rough picture of the ways in which the public engage in the conversation around the use of AI by LEAs. Social network analysis allows to identify the strong nodes and the patterns of relationships (Milory & Llamas, 2013) (Wasserman & Faus, 1994) by pinpointing which accounts are the most engaged in social media activity such as sharing information, resources, promoting the accounts or resources of collaborating agencies and responding to citizen accounts.

By analysing this activity, we can determine which accounts are 'influencers', accounts that play important roles in signposting social media users towards relevant resources. Discourse overview allows to have an insight on the content that is shared on social media.

3.1.1 Social listening

The social listening exercise was conducted using CommonCrawl, an open web repository containing 3.1 billion web pages (e.g. news media, activist websites, blogs), where each month's worth of data totals more than 300 terabytes. The data processing complied with General Data Protection Regulation (GDPR) and the research ethics principles and practices. An algorithm was employed to search through the database for strings of keywords that identify topics of interest to the popAI project. Initially, AI policing-related keywords were identified and scanned manually emerging from the controversy ecosystem mapping (T3.1) and the framework of the crowdsourcing platform (T3.3). Following the initial scan, the keywords were scanned via an automated Digital Dashboard that monitors CommonCrawl historical web data.

The keywords used corresponded to the five major topics that emerged from D3.1, namely Biometric Identification, Predictive Policing, Cyber Operations, Police Hacking Operations, Decision making in the justice systems. We also organised a list of subtopics indicating possible citizens' concerns: privacy, efficiency, reliability, accuracy, legitimacy, transparency, accountability, discrimination. For the social listening exercise, some of the subtopics were grouped together as they could be used interchangeably in online conversations. Therefore, for each topic we analysed the following set of subtopics(1) privacy, (2) efficiency, reliability, accuracy, (3) legitimacy, (4) transparency, accountability, (5) discrimination. Each subtopic was then further divided into two substrands - positive and negative[7]. In the positive substrand, keywords that could indicate a favourable disposition on the part of the communicator were included. An example of this would be the combination of "privacy" with keywords such as "acknowledge", "promote", "respect", "safeguard", "by design", "upholding". In the negative substrand, the keywords paired with "privacy" were "disrupt", "impedes'', "does not acknowledge", "not sensitive", "harm", "abuse", "restrict". We also developed a set of keywords to exclude to avoid ambiguity in our analysis As an example, in the positive substrand of "Legitimacy" keywords such as "no legitimacy" and "low legitimacy" were excluded[8].

Apart from the positive and negative strands, a neutral one was also included. This substrand contained only keywords that define the subtopic. For example, the keyword "privacy" was identified as neutral.

Besides the manual definition of the keywords' sentiment, results were run through a roBERTa-base model (Liu, et al., 2019), a sentiment analysis tool, trained on 58 million tweets - that scores the positive, negative, and neutral sentiment of each result. This model does not score in absolute terms a certain result can be classified as 70% positive, 20% neutral, and 10% negative.

[7] For a visual overview, please see Annex 1 "Topics, Subtopics, Substrands chart"
[8] For the full set of keywords used to define our search, please see Annex 2 "Social listening keywords".

Three test searches were firstly conducted on CommonCrawl and consequent refinements that the research team undertook to accommodate the multitude of goals, topics, and constraints of the analysis. Each test search encompasses a month worth of data on CommonCrawl, namely January 2020.

Initially, we set out to conduct the sentiment analysis limited to the keywords provided. The first test run clearly showed that there was a high number of false positives and false negatives. That is to say, there were results in the "positive" substrand that technically included a combination of the keywords, but the text sentiment was negative and vice versa. This led to the necessity to include a third "neutral" substrand for each subtopic, where no sentiment-catching keywords are used. At the end of the research, a sentiment analysis algorithm would be employed that would be fed with all hits from all substrands (positive, negative, and neutral) and score each one on negativity, positivity, and neutrality to have the broadest possible dataset. The averaging of the sentiment score for each subtopic, and the removal of duplicates, minimises influence on the results.

During the tests, the keywords came up in results that did not constitute a comprehensive text, but represented a string of words, most probably used in websites for Search Engine Optimisation, caches or other system files that were public on the internet. For this, the team had to employ another algorithm to sort through the data and determine if a certain result was, in fact, a useful 'signal' or 'noise' that had to be removed. To teach the algorithm to recognise the meaningful results, after each of the three test runs the team manually went through 250 results for each topic, subtopic and substrand, totalling approximately 18 000 hits. In total, more than 50 000 results were categorised as relevant or not and were fed into the algorithm until it succeeded in categorising correctly automatically.

The data extracted cannot be attributed to a specific geographical location, therefore the sentiment expressed cannot be indicative for the population of a country or continent. In the context of popAI, it means that the data could not be limited in a European context. Yet, internet users interact with data from all over the world and can be influenced regardless of the origin of a post, text, or news article. Furthermore, while CommonCrawl provides access to a vast amount of data, it still does not cover the whole internet and the research was limited to the keywords extracted by the controversies identified in previous tasks (specifically Task 3.1). This means that discourses internet users might have used to discuss certain topics could have been excluded.

3.1.2 Network analysis and discourse overview via NLP

For the current project, we focus on the Twitter social media platform. Twitter is a social media website that functions as a micro-blogging platform. The platform's use for microblogging and making announcements make it a highly dynamic environment where members of the public can receive updates from government agencies, politicians, official institutions, activist groups, and journalists, as well as interact with them and discuss with other users. By creating an account, Twitter users may post short text strings consisting of 280 characters or less called 'tweets'. While users can also post images, videos, and a few other types of media, here we focus only on text. Similarly to other social networking sites, Twitter allows users to follow other users, be followed, and interact

with other users tweets by 'liking' a tweet, 'retweeting' it (essentially re-posting a tweet to amplify its reach), mentioning other users, and replying to other users' tweets. Tweets can optionally include a hashtag, a short string of words following the # character. Hashtags are used to index and identify discussion topics on Twitter. Similarly, to the social listening, the hashtags as well as the handles that were used to conduct the analysis also emerged from the controversy ecosystem mapping task (T3.1).

To analyse data on Twitter, we used both network analysis and natural language processing techniques. Network analysis is used to study the interactions of official accounts with one another and the public via tweets, replies, and quotes. After collecting tweet data from the Twitter API, we constructed an interaction network where individual accounts were represented as nodes, and the interactions between them as connections. We then calculated various metrics for this network such as the 'centrality' and 'PageRank' of given nodes. PageRank and (Betweenness) Centrality are both metrics of influence in a network. PageRank is an algorithm used by google to rank the results of webpages returned by their search engine, on the assumption that pages with many incoming and outgoing links are more likely to be visited as they are both directed towards and themselves direct towards other relevant pages. Centrality is calculated by considering the shortest path between any pair of nodes and counting how many of these shortest paths pass through each node. Hence, nodes with a high centrality are those that are part of the shortest paths linking many other nodes. In both metrics, a node is considered influential because any journey on that network (for instance, a Twitter user navigating the site by viewing tweets and following links to the accounts that are mentioned) is likely to call at those influential nodes, giving those nodes an important role in signposting where to go next.

Twitter provides APIs (application programming interface) for requesting data from the servers. To use these APIs, a Twitter developer account associated to a Twitter user account must be created and a developer agreement has that outlines the acceptable uses of data from twitter[9]. A standard developer account is subject to limitations on the nature of the data that can be requested[10]. The standard developer API access allows a developer to make 900 requests for user objects per 15 minutes. User objects can contain information from many different 'fields'[11], but for the current work we request the following information: account creation date, account handle (unique user name) and screen name (non-unique name), user id number (unique and unchangeable number) , 'verified' status which indicates whether a given twitter account has been vetted for authenticity by Twitter (for public figures and official agencies, part of the rationale for this is to address the issue of potential imposters), number of followers and number followed accounts, number of tweets and number of 'likes' given, optional account bio/description (free text entry with 160 character limit), optional external URL, optional location (free text entry with 30 character limit).

The standard API allows for between 450 and 1500 tweet object requests per 15-minute period depending on whether the search request is a general search or a search for the tweet timelines of

[9] https://developer.twitter.com/en/developer-terms/agreement-and-policy
[10] https://developer.twitter.com/en/docs/twitter-api/getting-started/about-twitter-api
[11] https://developer.twitter.com/en/docs/twitter-api/data-dictionary/object-model/user

individual users. Tweet objects can again contain information from a variety of different fields[12]. In the current work we request the following information: text of the tweet, author id number, list of mentioned accounts, list of included hashtags, number of retweets, likes and replies.

While both tweets and accounts can contain optional geolocation data, only 1-3% of tweets have this functionality enabled (Zohar, 2021), and so restricting our analysis to only Europe is not feasible. When we consider the tweets of accounts relating to European agencies or institutions, we can be reasonably certain that the content of their tweets is relevant to the European context. However, when requesting data from the general search API, there is no way to ensure that the tweets objects we get are all from accounts based in Europe.

For identifying influential accounts on twitter, we consider the local network surrounding a list of accounts emerging from the controversy ecosystem mapping task (T3.1) as listed in Table 1. Rather than considering the network formed by Twitter's 'follow' relationship, we consider the network formed by interactions between users (mentions, replies, and retweets). The standard API limitations mean that general search queries only return results from the past 7 days. When requesting the timeline of an individual account, we can retrieve up to 3200 of the account's most recent tweets. Depending on how active an account is, those 3200 tweets could cover a timeframe of weeks, months or even years. Owing to these limitations, the interaction network we can construct is only a snapshot of the network as it is at the time we collect the data and might be different if we were to search the full Twitter archive or collect data via the same methods at a different time. The network analysis we carry out is intended to illustrate the dynamics and structure of social media interaction at a moment in time.

Name	Handle
Access Now	accessnow
ARTICLE 19 Campaigns | #Act4Expression	Act4Expression
Algorithmic Justice League	AJLUnited
Algorithmic societies	AlgorithmicSoc
AlgorithmWatch	algorithmwatch
Amnesty International	amnesty
Amnesty Tech	AmnestyTech
ARTICLE 19	article19org
Andrew Stroehlein	astroehlein
Border Violence Monitoring Network	Border_Violence
DataEthics	DataEthicsEU
Data Justice Lab	DataJusticeLab
EDRi	edri
ECNL	enablingNGOlaw
Fair Trials	fairtrials
Freedom not Fear	fnf_eu
Hermes Center	HermesCenter

[12] https://developer.twitter.com/en/docs/twitter-api/data-dictionary/object-model/tweet

Homo Digitalis	Homo_Digitalis_
Human Rights Watch	hrw
Tactical Tech	Info_Activism
IPVM	ipvideo
Kenneth Roth	KenRoth
Lajla Fetic	lajlafetic
Migration and Technology Monitor	migration_tech
Ministry of Privacy	ministryprivacy
Privacy International	privacyint
Privacy Matters	PrivacyMatters
Reclaim Your Face	ReclaimYourFace
Statewatch	StatewatchEU
Turin Privacy Group	TurinPrivacy
Margrethe Vestager	vestager
Institute for Ethics in Artificial Intelligence	IEAITUM
Artificial Intelligence Lab Brussels	aibrussels
OEIAC	OEIAC_UdG
FIRE Forum	FIRE_ForumEU
DG DEFIS #StrongerTogether	defis_eu
NextEOS	next_eos
European Defence Agency	EUDefenceAgency
Frontex	Frontex
EOS	EOS_EU
TU MÃ¼nchen	TU_Muenchen
Elecnor Deimos	ElecnorDeimos
Fraunhofer IAIS	FraunhoferIAIS
ARMINES	_ARMINES_
Secure-IC	SecureIC
Digital EU	DigitalEU
Efus	Efusnews
EU Institute for Security Studies	EU_ISS
SIPRI	SIPRIorg
IIT	IITalk
AI 4Copernicus	AI4Copernicus
AI-on-Demand Platform	AI4EU
DIH4AI project	dih4ai
I-NERGY	inergy_h2020
AIPlan4EU	AIPlan4EU
CLAIRE	vision_claire
TAILOR EU Network	eu_tailor
HumaneAI	ai_humane
AI4Media	ai4mediaproject
OECD Innovation	OECDinnovation
CyberSec_EU	Cybersec_EU

CISPA	CISPA
NOTIONES_EU	NOTIONES_EU
Fraunhofer Presse	Fraunhofer
CBRNE Ltd	CBRNE_Ltd
FOI	FOIresearch
KU Leuven	KU_Leuven
Ertzaintza	ertzaintzaEJGV
Bayerisches Innenministerium	BayStMI
CEA_Officiel	CEA_Officiel
AIT	AITtomorrow2day
BMI	BMI_OE
Bundesbereitschaftspolizei	bpol_bepo
CNRS	CNRS
DFKI	DFKI
ENGINEERING	EngineeringSpa
EKETA-CERTH	CERTHellas
Europol	Europol
Vicomtech	Vicomtech
Hellenic Police	hellenicpolice
Herta	hertasecurity
INOV	INOVinesc
LINKS Foundation	LinksFoundation
Ministere de l'Interieur et des Outre-mer	Interieur_Gouv
Ministerio del Interior	interiorgob
Il Viminale	Viminale
Pluribus One	pluribus_one
Police Federale	policefederale
Policie ÄŒR	PolicieCZ
Politsei	Politsei
Sheffield Hallam University	sheffhallamuni
TNO Research	TNO_Research
VTT	VTTFinland
Thales Group	thalesgroup
VoiceInteraction	v_interaction
Politecnica Madrid	La_UPM
Web-IQ	webiqnl
PolBru	zpz_polbru

Table 1 Hashtags of stakeholder accounts

To construct the network, we request the timelines of the twitter accounts listed in Table 1 and analyse the content of their tweets, as well as examine which accounts they interact with via replies, mentions or quote-tweets. We then further request the twitter timelines of those mentioned accounts to build a larger picture of the local interaction network. We can then use network metrics

such as PageRank and centrality to determine which accounts are influential in the network. PageRank and centrality are metrics that determine the importance of objects in a network, on the assumption that important nodes in a network are more likely to have many connections. In the context of a social interaction network, nodes with high pagerank and centrality play a role as a 'hub' from which information can spread and flow through. To collect data for this, we collect the recent timelines (3200 most recent Tweets) of the accounts listed in Table 1, and then search for replies and quotes of other accounts. We then perform a secondary search of the timelines of those accounts they have mentioned or replied to. With this corpus of tweets, we then build a network of accounts, with two accounts connected if their twitter timelines show interaction with one another.

In addition to this, we also use the general search API to search for both tweets that are directed at these accounts, as well as the discussion on the hashtags listed in Table 2 to get an overview of both general discourse on twitter around the hashtag topics, as well as the comments directed at official institutions, agencies, and stakeholders. While the intended use of hashtags is to index topics, users can include unrelated hashtags in a tweet meaning that using hashtags as a method of collecting data can result in collecting unrelated tweets and this should be taken in account when analysing data collected via querying the API for hashtags. To get an overview of the social media discourse, we collect data from the Twitter API by querying the listed hashtags, and querying for Tweets directed at the handles listed in Table 1. We then use a natural language processing (NLP) library in Python called Spacy, to load a pre-trained language model that can 'tokenise' and 'tag' the words in the tweets, that is, identifying the unique words as nouns, verbs, adjectives, allowing us to count their frequency. An advantage of this approach is that it allows us to identify unique words more robustly than merely comparing the text strings. For example, different verb conjugations such as 'walk', 'walking' and 'walked' can all be identified as instances of the same root word but in different contexts, where merely considering the characters strings would count them as three different words. However, a significant limitation is the mixed languages we collected data in. While Spacy does have models for several languages, just under 70% of the data we collected is in English, and without a robust or reliable way to perform language detection, we cannot select appropriate models for the remaining data. The spacy library can still parse and tokenise non-English data, but it is not able to identify non-English nouns, verbs, or adjectives.

BanBiometricSurveillance
ReclaimyourFace
privacy
DigitalRights
MassSurveillance
BiometricMassSurveillance
AIAct
HumanRights

BanFacialRecognition
AIact
surveillance
facialrecognition
BanTheScan
BanBS
AI
biometrics
biometric
FaceSurveillance
banBS
algorithm
DataAct
ArtificialIntelligence
ClearviewAI
iBorderCtrl
Palantir
dataethics
aiethics
BigData
DataEthics
cybersecurity
cyberoperations
cyberwar
keycrime
predictivePolicing
Precobs
GangsMatrix
CrimeAnticipationSystem
IPolice

Table 2 Hashtags containing keywords, phrases names of companies and names of technologies

To attempt to provide a point of comparison, we also considered the content aggregation and discussion site Reddit. On Reddit, users can post links, images or pure text which is then rated and responded to by other reddit users, showing popular content to at the top of the site while less active (and hence presumably less relevant) material is moved off the front page of topics. Reddit is organised into groups known as subreddits which typically focus on particular topics, which can be broadly or narrowly defined. We gathered posts from two popular subreddits, 'r/Europe' and 'r/Technology' to try and find discussions about technology relevant to the European context. To do this, we used the free API of an open-source archive of Reddit posts called PushShift. A limitation of this approach is that reddit posts cannot be geolocated at all, and so on the r/Technology subreddit, narrowing our search to only return discussions relevant to the European context is not feasible. Alternatively, on the r/Europe subreddit, discussions pertaining to technology can be located by searching for the keywords that comprise the hashtags list of Table Y. However, we found that discourse about technology was typically a lot more rarefied on this subreddit, with most users discussing more broad geopolitical topics.

3.2 Findings

3.2.1 Social listening

The findings of the analysis conducted on the CommonCrawl database will be presented in D3.3 "Citizen produced priorities and recommendations for addressing AI in the security domain" (March 2023).

3.2.2 Social Network Analysis on Twitter

From each of the listed accounts, Twitter timelines were collected consisting of the most recent 3200 tweets (or fewer where the account has less than 3200 tweets) to capture the recent interaction activity of these accounts. Following this, we gathered the most recent 1000 tweets of all accounts that were mentioned in these timelines and checked to see if they interacted with any of the other accounts. By following this methodology, we discovered 1457 unique users engaging in 51980 unique interactions. Accounts which only interact with a single other account were removed for visualisation purposes as they do not count as influential accounts in their own right, leaving a network consisting of 969 nodes representing accounts and 6954 edges representing interactions between them. 279 accounts were 'verified' accounts, typically belonging to public figures, journalists or the accounts of companies and institutions. Amongst the accounts in our network, only 1.6% of the tweets we collected had geolocation enabled or have entered a meaningful location as part of their profile information.

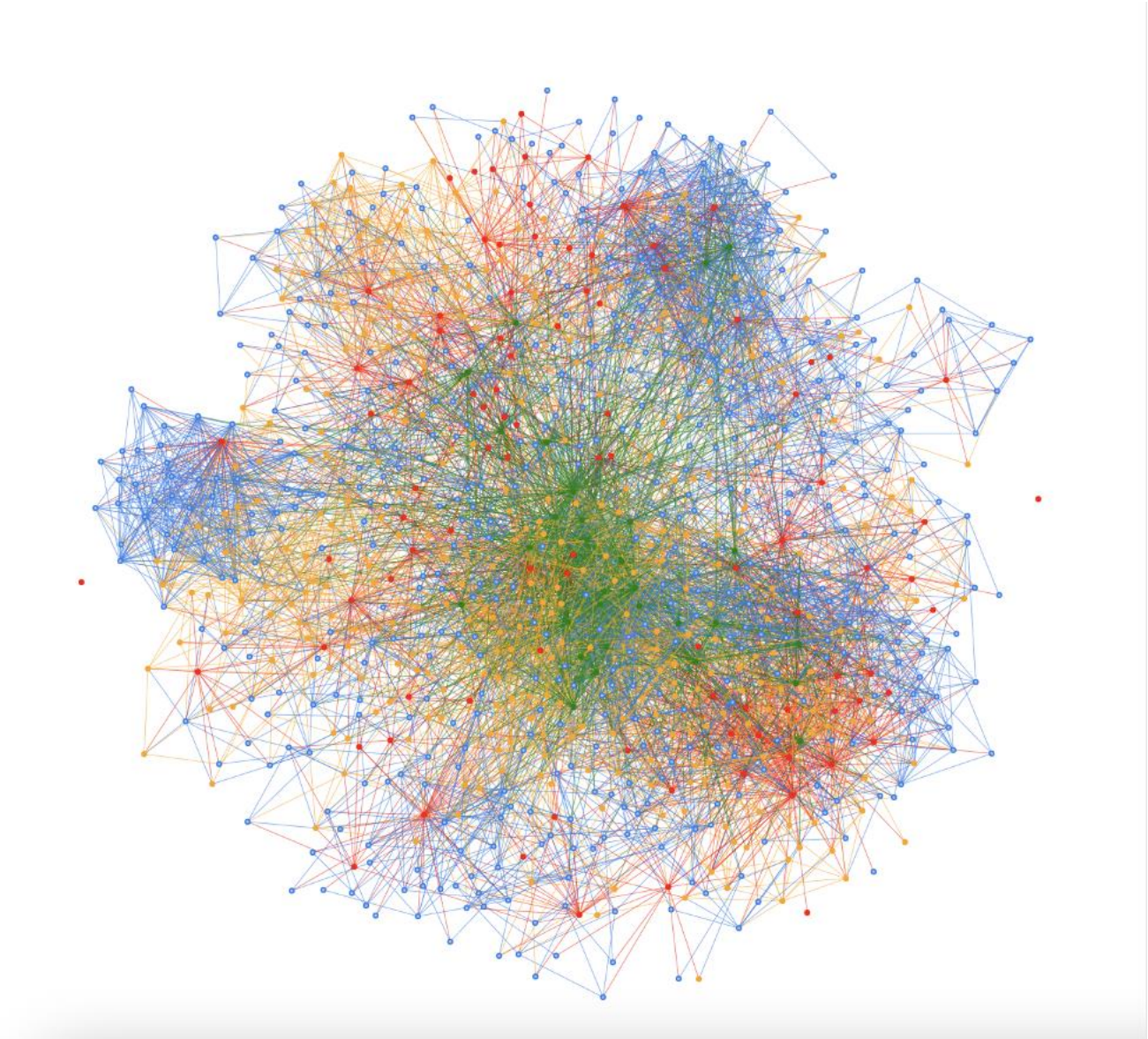


Figure 1 Snapshot interaction network showing the interactions between listed account handles (red), verified accounts (orange), the 50 most influential accounts (that were not already listed or verified) as determined by PageRank(green) and unverified/civilian accounts (blue)

Figure 1 shows a visualisation of the local network surrounding the accounts listed in Table 1, with colours indicating whether they are listed in Table 1 (red), verified accounts (orange), or unverified civilian accounts (blue). The 50 most influential accounts, as determined by the top 50 highest PageRank and Centrality scores are coloured red unless they are already listed in Table 1. It is important to note that distance and closeness on this visualisation are not necessarily significant. Linked nodes tend to be pulled closer together, but the distances between nodes carry no significance in and of themselves. The PageRank and Centrality scores of some of these significant accounts are given in Tables 3 and 4. Table 3 lists the accounts which were previously given in Table 1, and Table 4 contains accounts which were determined to be of importance by our analysis but were not already listed.

Handle	Pagerank	Centrality
@Europol	0.00536	0.08682
@accessnow	0.00443	0.07949
@SIPRIorg	0.00391	0.05020
@privacyint	0.00385	0.06171
@Frontex	0.00350	0.06066
@hrw	0.00347	0.06171
@edri	0.00340	0.06380
@La_UPM	0.00339	0.05962
@Fraunhofer	0.00332	0.05753
@bpol_bepo	0.00322	0.04707
@algorithmwatch	0.00320	0.05334
@article19org	0.00316	0.04916
@amnesty	0.00301	0.05230
@thalesgroup	0.00292	0.03870
@vestager	0.00290	0.04916

Table 3 PageRank and Centrality of accounts that were previously listed in Table 1

Handle	Pagerank	Centrality
EU_Commission	0.01072	0.19142
YouTube	0.00741	0.11087
UN	0.00680	0.11715
Europarl_EN	0.00431	0.07949
vonderleyen	0.00414	0.07740
YlvaJohansson	0.00406	0.07426
BMBF_Bund	0.00403	0.07217

WHO	0.00356	0.05857
nytimes	0.00351	0.05439
ERC_Research	0.00333	0.05334
esa	0.00325	0.05230
EUCouncil	0.00304	0.05648
BMI_Bund	0.00293	0.04916
EmmanuelMacron	0.00287	0.04811
ellajakubowska1	0.00283	0.05230

Table 4 The PageRank and centrality of influential accounts that were not already listed in Table 1

3.2.3 Discourse overview via NLP tagging

After collecting tweets directed at the stakeholder accounts, as well as tweets on the hashtags listed in Table 2, we used a pre-trained English language model from the Spacy python library to tokenise tweet text into words, and then tagged those words as nouns, verbs, and adjectives. Figure 2 shows a word cloud of the most commonly occurring words on the hashtags from Table 2, with more common words occupying more space. Figure 3 shows a similar cloud for the words in tweets directed to the stakeholder accounts in Table 1.

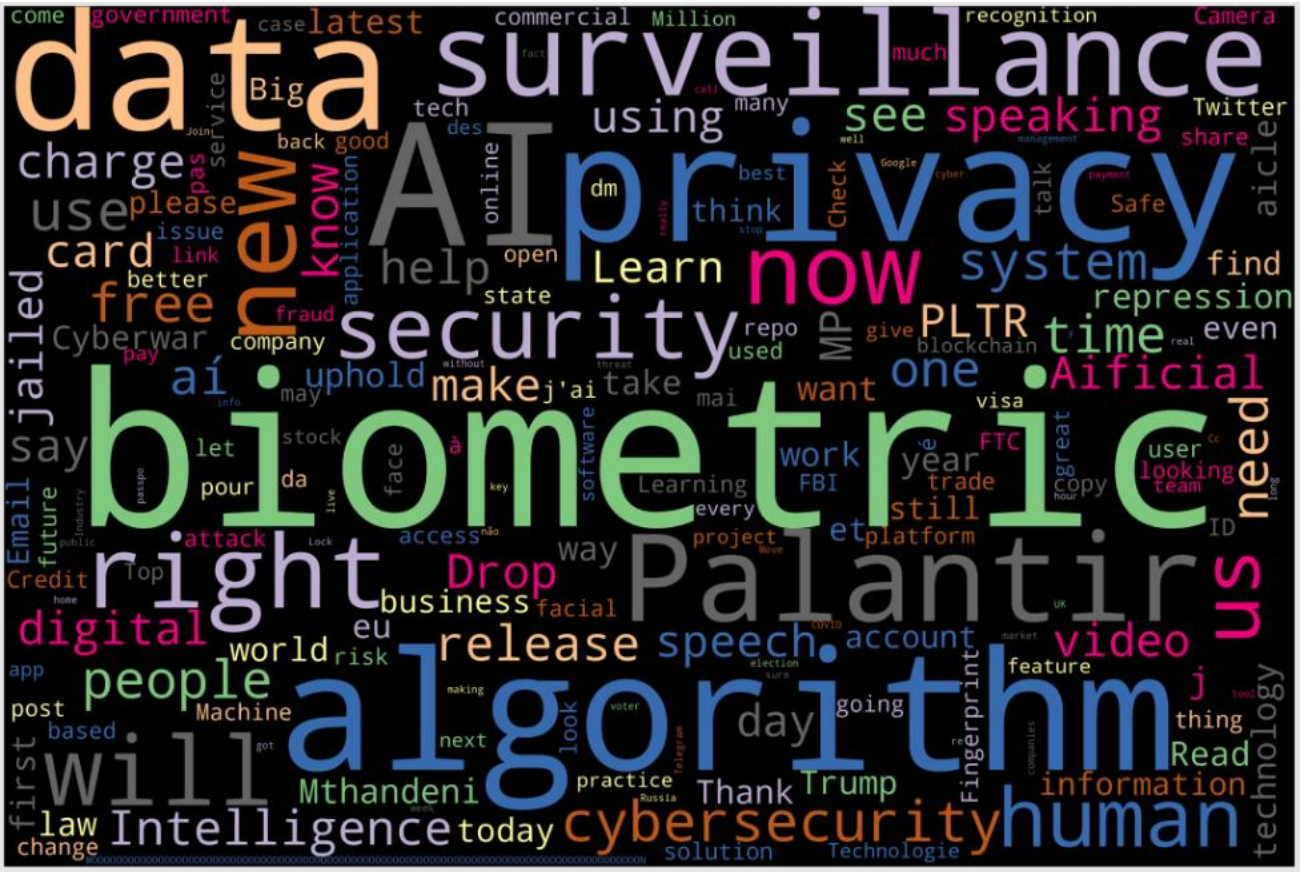


Figure 2 A word cloud of terms used in discussions occurring on the listed hashtags

Figure 3 A word cloud of terms used in tweets directed at the listed accounts from members of the public

Amongst tweets collected via querying the hashtags, the Spacy NLP English language model identifies the most commonly used verbs, adjectives and adverbs as:

- verbs: need, know, talk, help, walk, use, think, learn, want and read
- adjectives: biometric, new, human, free, late, good, big, bad, safe and digital
- adverbs: maybe, soon, long, actually, right, far, probably, ago, well and anonymously

We also considered word pairings, so-called 'bigrams'. Many of these were common words combined with articles of grammar, however some recurring phrases included: '**human rights**', '**free speech**' and '**against repression**'.

Amongst tweets directed at the accounts on our list, the most used verbs, adjectives, and adverbs are:

- verbs: 'stop', 'hold', 'commit', 'fund', 'cooperate', 'expect', 'issue','have','fall' and 'send'.
- adjectives: 'accountable', 'bad', 'russian', 'human', 'shameful', 'ukrainian', 'fair','peaceful', 'absurd' and 'civilized'.
- adverbs: democratically, clearly, maybe, right, long, actually, anymore, well, late and soon

Salient bigrams tweeted at the accounts include: '**held accountable**', and '**human rights**'.

Our searches on Reddit were not able to gain as robust an insight as we lacked a significant amount of data. Many posts relating to these topics involved posting links to news stories with text quoted verbatim from the articles. Due to the nature of the platform and international demographics discussion was generally US-centric and conducted in the English language, even when discussing controversies in non-English speaking countries. We struggled to draw out observations from the data we gathered as the discourse did not follow predictable patterns of speech, mostly consisting of chatter between users reacting to one another's statements and remarks that require the rest of the discussion thread as context to fully parse. Comments that either repeat verbatim earlier comments or quote from an article are also common.

However, we noted the following:

- Discussion is chiefly focused on comparisons to popular culture - references to Skynet and Terminator were frequent.
- 'West vs East' comparisons were also common, particularly with negative discussion and references to China and its 'credit score' system.
- From inspection of some text posts, we noted that the impact of surveillance technologies on women and minorities were a common theme, with typically very low trust of corporate and government entities in this context.
- Discussion of surveillance technologies in the context of capitalism was common, i.e. discussing profit generated from selling data and implementing AI was a common theme discussion rather than the social utility of said technologies.
- Regarding specific technologies such as the use of AI in creating 'deep fake' images and videos, or the deployment of AI to detect such fakes, we saw users expressing feelings of both interest in the technology as well as generalised anxiety over the idea that such technology might be misused, with no specific instances of misuse being mentioned or discussed.

4 Discussion on the Findings

The interaction network we observed on Twitter in Figure 1. shows that the accounts that emerged from the controversy ecosystem mapping (Table 1) are well networked with one another and well networked with 'verified' accounts, with 29% of the overall network being verified accounts. The metrics of PageRank and centrality both concord on which accounts are considered influential, ranking the top 15 accounts in the same way. We observe that the stakeholder accounts are both well networked with the most influential accounts as well as being influential in the network in their own right. Of the accounts we observe in this network, civilian and activist group accounts do not appear to hold influential roles, with the most influential accounts belonging to established organisations and political figures. In order to address this, an effective strategy for amplifying the voices of civilian and activist groups would be for influential accounts to retweet these groups, or quote-tweet them and signpost them to their own following. In particular, we note that the most influential nodes in the network are the official twitter accounts of the EU Commission, European Parliament and the United Nations. These accounts might hold influential roles in the network as a result of their broader functions rather than any specific influence in the specific fields of technology and security. We also note that the account belonging to the streaming service YouTube has an influential role in the network, probably owing to the role of YouTube in hosting video content which

is used by many government agencies to communicate to the public, as well as journalistic publications and journalists themselves to share video content. However, another reason for the appearance of YouTube in this network is civilian discussion of the algorithms that determine content delivery and recommendation to users of the platform. Many references to the 'Youtube algorithm' were made, sharing speculation about the algorithms used to suggest and recommend content, with a mixture of praise for the results whilst expressing concern about the lack of transparency in such algorithmic approaches.

Discussion on Twitter tends to consist mostly of short sentences with impactful phrases, likely owing to the character limits of tweets and the difficulty of threading or parsing longer form discussion. As we can see in Figures 2 and 3, much of the discourse referred to current events around the time of the data collection, with Russia and Ukraine featuring prominently in general twitter discourse, but even more so in discussions and comments directed at European institutions and agencies. This is likely owing to the nature or perception of Twitter as a platform to discuss and comment on contemporary events. In the context of discussions around biometric identification, comments we detected focused on the privacy of individuals data and the risk of this data being stolen and/or misused. Specifically, there were concerns about the use of AI technologies to create fake biometric data, such as using AI methods to record and then generate fake voice clips so to pass voice authentication or otherwise commit identity fraud.

5 Conclusions

This deliverable explored the public discourses around the employment of AI in the sensitive area of law enforcement, focusing particularly on citizens' discourses around the emerging controversies on this topic. Technological advancements have always been used to support crime investigation and even crime prediction. It is of no surprise then the ever-increasing use of AI systems in the field of law enforcement. The employment of AI in the civil security domain is broadly accepted and endorsed at the EU level (EC, 2019) while raising afresh concerns regarding the potential impact and harm to the society and fundamental rights.

AI has a wide range of functionalities and applications in the security domain[13]. The main topics that have been identified as the major ones in the EU are predictive policing, facial recognition, AI and criminal justice, and AI and borders (Gonzalez Fuster, 2020). Indeed, the deployment of AI systems based on facial recognition have raised great concerns in the EU and revoked solid initiatives such as campaigns and demonstrations. The legal taxonomy report[14] also highlighted how regulatory documents that aim to govern AI are greatly focused on facial recognition, confirming the importance

[13] D3.1 followed a more detailed taxonomy of the civil security domain structure to investigate controversial cases of AI use namely, Crime prevention, Crime Investigation, Migration, Asylum, and Border control, Administration of Justice, Cyber Operations for law enforcement, and Training. See also D2.1 Functionality taxonomy and emerging practices and trends, submitted in April 2022

[14] D2.2 Legal; casework taxonomy: emerging trends and scenarios

of potential risks emerging from this technology. This report supports these claims by showing that biometrics are at the centre of the Twitter debate.

The insights of discussions on Reddit around biometric identification detected another concern in addition to privacy and data protection, that of data misuse. In particular, the risk of data misuse concerned the use of AI to create fake biometric data for identity fraud.

Most AI systems in the security domain are surveillance-based and this can partly explain the societal mistrust and scepticism. Studies indicate low levels of trust in AI systems especially in high-income European countries (e.g., Italy, Germany, France, Belgium, Sweden, the Netherlands) (WEF, 2022). Analysis on Twitter data indicated the increased concerns over the impact of surveillance technologies on women and minorities indicating very low trust to corporate and government entities. Indeed, analysis on reddit indicated that the discussion of surveillance technologies was commonly in the context of capitalism rather than the societal benefits and as a response to societal needs. Civil society organisations, as well as EU bodies promote actions and initiatives to increase transparency which could foster trust. To this end, discourse analysis on Twitter indicated that recurring phrases were 'human rights', 'free speech' and 'against repression' while salient topics emerged include accountability and human rights.

However, network analysis showed that civilian and activist group accounts do not appear to hold influential roles which might result in their campaigns and actions not reaching the wider society. This is in accordance with the findings of ecosystem mapping (D3.1) indicating that while civil organisations were active in revealing emergent risks from the use of AI for law enforcement, they were underrepresented in research projects in security domain.

The results of social media listening to be presented in the report 3.4 (due in March 2023) will shed light on the sentiment with which AI use in specific LEAs' areas are discussed. The findings will explain further what the tendencies in the different discourses are, positive, negative, or neutral.

This report focused on the identification of the network that discusses the use of AI in the civil security domain and the influencers regarding the public discourse on social media and specifically twitter. The identification of the most influential accounts will support the popAI future research and importantly the dissemination and engagement with the project's pandect of recommendations for the ethical use of AI for LEAs.

6 References

Akhgar, B. et al., 2022. *Accountability Principles for Artificial Intelligence (AP4AI) in the Internal Security Domain,* s.l.: Europol Innovation Lab; CENTRIC.

Cohen, J., 2017. Surveillance vs. privacy: effects and implications. In: D. Gray & S. Henderson, a cura di *Cambridge Handbook of Surveillance Law.* New York: Cambridge University Press, pp. 455-469.

Drobotowicz, K., Kauppinen, M. & Kujala, S., 2021. Trustworthy AI Services in the Public Sector: What Are Citizens Saying about It?. *International Working COnference on Requirements Engineering*

EC, 2022. *European Group on Ethics in Science and New Technologies.* [Online] Available at: https://research-and-innovation.ec.europa.eu/strategy/support-policy-making/scientific-support-eu-policies/ege_en

EC, 2019. *Ethics guidelines for trustworthy AI.* [Online] Available at: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai EC, 2019. *Building Trust in Human-Centric Artificial Intelligence,* Brussels: COM(2019) 168 final.

EC, 2016. *Stronger and Smarter Information Systems for Borders and Security,* Brussels: COM(2016) 205 final.

EC, 2014. *Ethics of security and surveillance technologies.* [Online] Available at: https://op.europa.eu/en/publication-detail/-/publication/6f1b3ce0-2810-4926-b185-54fc3225c969

EC, 2014. *Towards a thriving data-driven economy,* Brussels: COM(2014) 442 final.

EDRi, et al., 2021. *Prohibit predictive policing and profiling AI systems in law enforcement and criminal justice,* s.l.: Fair Trials, European Digital Rights (EDRi), Access Now, Algorithm Watch, European Disability Forum, European Not for Profit Law Centre, Panoptykon Foundation.

EDRi, et al., 2021. *Prohibit all Remote Biometric Identification in publicly accessible spaces,* s.l.: European Digital Rights, Access Now, ARTICLE19, Bits of Freedom, Chaos Computer Club, etc..

EDRi, AN, BoF & etc, 2021. *Prohibit emotion recognition in the Artificial Intelligence Act,* s.l.: Access Now, European Digital Rights, Bits of Freedom, ARTICLE19, ITPol.

EP, 2021. *European Parliament resolution of 6 October 2021 on artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters,* Strasbourg: EP.

Esposti, S. & Gomez, E., 2015. Acceptable surveillance-oriented Security Technologies: Insights from the SurPRISE Project. *Surveillance & Society,* 13(3/4), pp. 437-454.

Franko Aas, K., 2005. *Sentencing in the age of information: from Faust to Macintosh.* New York: Routledge.

Fuerstenau, M., 2017. Germany's facial recognition pilot program divides public. *DW.*

Gonzalez Fuster, G., 2020. *Artificial Intelligence and Law Enforcement - Impact on Fundamental Rights,* Brussels: European Parliament.

Harcourt, B., 2008. *Against Prediction - Profiling, Policing, and Punishing in an Acturial Age.* s.l.:The University of Chicago Press.

HLEG, 2019. *Ethics Guidelines for Trustworthy AI,* s.l.: EC.

Liu, Y., Ott, M. & Goyal, N., 2019. *RoBERTa: A Robustly Optimized BERT Pretraining Approach.* s.l.:s.n.

Lyon, D., 2003. *Surveillance as Social Sorting: Privacy, Risk, and Digital Discrimination.* London: Routledge.

Lyon, D., 2003. Surveillance after September 11. *Polity.*

Milory, L. & Llamas, C., 2013. *The handbook of language variation and change.* s.l.:Social networks.

Nissenbaum, H., 2009. *Privacy in Context: Technology, Policy, and the Integrity of Social Life.* Stanford: Stanford Law Books.

Ronsin, X. & Lampos, V., 2018. *In-depths study on the use of AI in judicial systems, notably AI applications processing judicial decisions and data,* Strasbourg: CEPEJ.

SURVEILLE, 2020. *SURVEILLE.* [Online] Available at: https://surveille.eui.eu

WEF, 2022. *5 charts that show what people around the world think about AI.* [Online] Available at: https://www.weforum.org/agenda/2022/01/artificial-intelligence-ai-technology-trust-survey

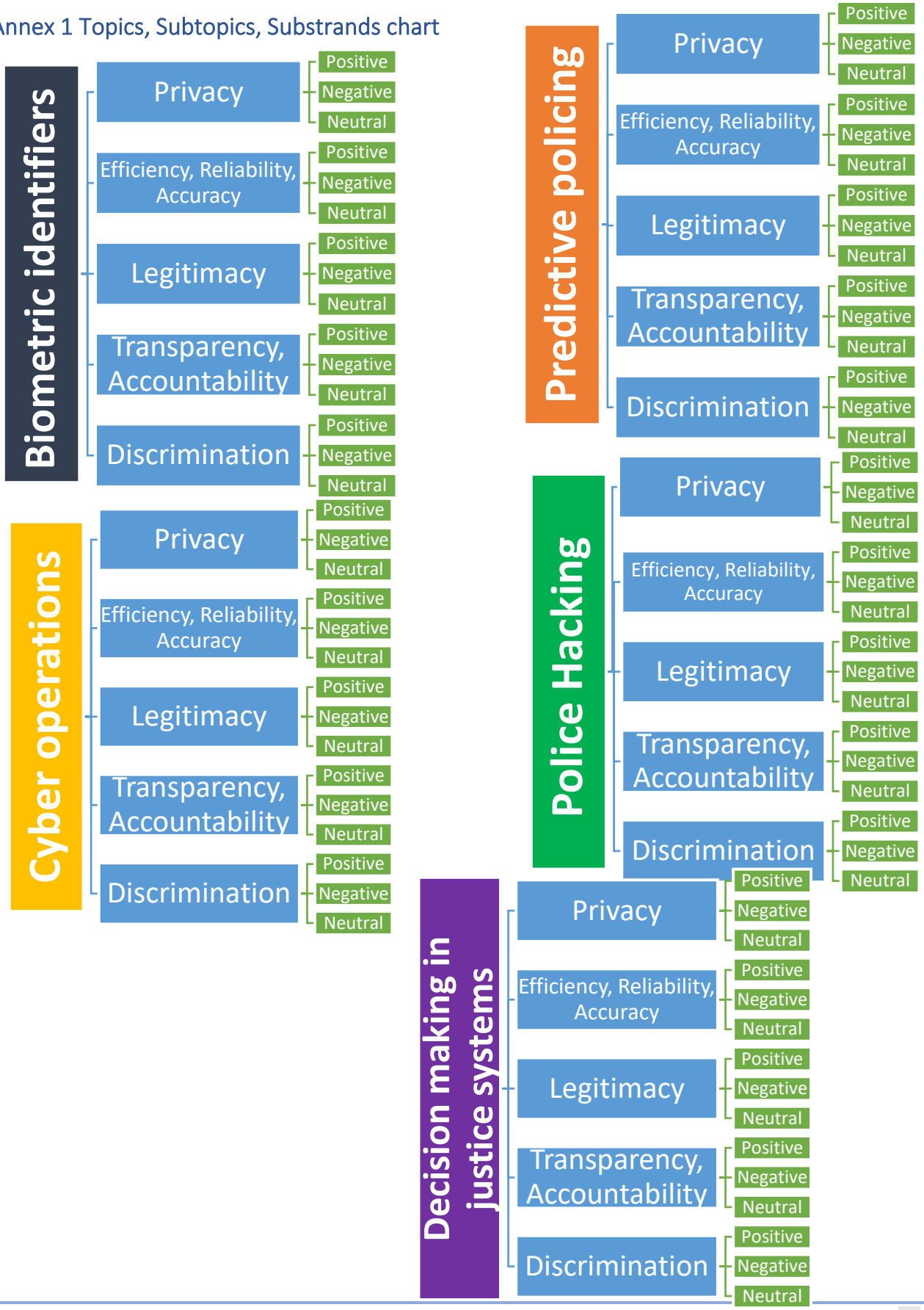
Wasserman, S. & Faus, K., 1994. *Social network analysis: Methods and applications.* s.l.:s.n.

Wilson, D., 2017. Algorithmic patrol: The futures of predictive policing. In: A. Zavrsnik, a cura di *Big Data, Crime and Social Control.* London and New York: Routledge, pp. 108-127.

Zohar, M., 2021. Geolocating tweets via spatial inspection of information inferred from tweet metafields. *International Journal of Applied Earth Observation and Geoinformation.*

7 ANNEXES

Annex 1 Topics, Subtopics, Substrands chart



Annex 2 Social listening keywords

Biometric identifiers

1. Privacy – negative
    - Keywords position 1: biometric identification | facial recognition | biometrics | biometric* authentication | fingerprint mapping | retina scan* | iris scan*
    - Keywords position 2: police | law enforcement | security agencies | enforcement agencies
    - Keywords Position 3: disrupts|impedes|does not acknowledge|not sensitive|harm*|abuse|restrict
    - Keywords Position 4: privacy
    - Exclusions: -Analytica - Autonomous vehicle - CBD Hemp

2. Privacy – positive
    - Keywords position 1: biometric identification | facial recognition | biometrics | biometric* authentication | fingerprint mapping | retina scan* | iris scan*
    - Keywords position 2: police | law enforcement | security agencies | enforcement agencies
    - Keywords position 3: acknowledge | promote | respect | safeguard | by design | upholding
    - Keywords position 4: privacy
    - Exclusions: -no safeguard -lacks safeguards -doesn't promote privacy -does not respect -doesn't respect -protectionCloud
3. Privacy
    - Keywords position 1: biometric identification | facial recognition | biometrics | biometric* authentication | fingerprint mapping | retina scan* | iris scan*
    - Keywords position 2: police | law enforcement | security agencies | enforcement agencies
    - Keywords Position 3: privacy
    - Exclusions: -Analytica - Autonomous vehicle - CBD Hemp

4. Efficiency, Reliability, Accuracy – positive
    - Keywords position 1: biometric identification | facial recognition | biometrics | biometric* authentication | fingerprint mapping | retina scan* | iris scan*
    - Keywords position 2: police | law enforcement | security agencies | enforcement agencies
    - Keywords Position 3: effectively | efficient | reliable | accurately | accurate | fight crime | prevent crime | increased security | support crime preventionKeywords
    - Exclusions: -low efficiency -low reliability -not reliable -not efficient -not accurate

5. Efficiency, Reliability, Accuracy – negative
    - Keywords position 1: biometric identification | facial recognition | biometrics | biometric* authentication | fingerprint mapping | retina scan* | iris scan*
    - Keywords position 2: police | law enforcement | security agencies | enforcement agencies
    - Keywords Position 3: ineffective | inefficient | not reliable | not accurate | inaccurate | unreliable

6. Efficiency, Reliability, Accuracy

- Keywords position 1: biometric identification | facial recognition | biometrics | biometric* authentication | fingerprint mapping | retina scan* | iris scan*
- Keywords position 2: police | law enforcement | security agencies | enforcement agencies
- Keywords Position 3: efficiency | reliability | accuracy | efficient | reliable | accurate
- Exclusions:

7. Legitimacy – positive
    - Keywords position 1: biometric identification | facial recognition | biometrics | biometric* authentication | fingerprint mapping | retina scan* | iris scan*
    - Keywords position 2: police | law enforcement | security agencies | enforcement agencies
    - Keywords Position 3: legitimacy | legitimate | is legitimate
    - Exclusions: -not legitimate -low legitimacy

8. Legitimacy – negative
    - Keywords position 1: biometric identification | facial recognition | biometrics | biometric* authentication | fingerprint mapping | retina scan* | iris scan*
    - Keywords position 2: police | law enforcement | security agencies | enforcement agencies
    - Keywords Position 3: not legitimate | illegal

9. Legitimacy
    - Keywords position 1: biometric identification | facial recognition | biometrics | biometric* authentication | fingerprint mapping | retina scan* | iris scan*
    - Keywords position 2: police | law enforcement | security agencies | enforcement agencies
    - Keywords Position 3: legitimacy | legitimate | lawfulness | lawful

10. Transparency, Accountability – positive
    - Keywords position 1: biometric identification | facial recognition | biometrics | biometric* authentication | fingerprint mapping | retina scan* | iris scan*
    - Keywords position 2: police | law enforcement | security agencies | enforcement agencies
    - Keywords Position 3: disrupts|impedes|does not acknowledge|not sensitive|harm*|abuse|restrict
    - Exclusions: -low transparency -no transparency -no accountability -no oversight -lack of oversight -lack of accountability

11. Transparency, Accountability – negative
    - Keywords position 1: biometric identification | facial recognition | biometrics | biometric* authentication | fingerprint mapping | retina scan* | iris scan*
    - Keywords position 2: police | law enforcement | security agencies | enforcement agencies
    - Keywords Position 3: no | lack | deficient | weak | ineffective | lack of
    - Keywords Position 4: transparent | accountable | transparency | accountability | oversight

12. Transparency, Accountability
    - Keywords position 1: biometric identification | facial recognition | biometrics | biometric* authentication | fingerprint mapping | retina scan* | iris scan*
    - Keywords position 2: police | law enforcement | security agencies | enforcement agencies
    - Keywords Position 3: transparent | accountable | transparency | accountability | oversight

13. Discrimination – positive
    - Keywords position 1: biometric identification | facial recognition | biometrics | biometric* authentication | fingerprint mapping | retina scan* | iris scan*
    - Keywords position 2: police | law enforcement | security agencies | enforcement agencies
    - Keywords Position 3: no discrimination | impartial | do not discriminate

14. Discrimination – negative
    - Keywords position 1: biometric identification | facial recognition | biometrics | biometric* authentication | fingerprint mapping | retina scan* | iris scan*
    - Keywords position 2: police | law enforcement | security agencies | enforcement agencies
    - Keywords Position 3: promote discrimination | promote racism | racial bias | disproportionately | inequality
    - Exclusions: -CBD Hemp -metersBeyond

15. Discrimination
    - Keywords position 1: biometric identification | facial recognition | biometrics | biometric* authentication | fingerprint mapping | retina scan* | iris scan*
    - Keywords position 2: police | law enforcement | security agencies | enforcement agencies
    - Keywords Position 3: discrimination | impartial | bias | discriminate | race | minorities

Predictive policing

1. Privacy – negative
   - Keywords position 1: predictive policing | preventive policing | crime forecasting
   - Keywords position 2: disrupts|impedes|does not acknowledge|not sensitive|harm*|abuse|restrict
   - Keywords Position 3: privacy
   - Exclusions: -seekersBaiduBenefits

2. Privacy – positive
   - Keywords position 1: predictive policing | preventive policing | crime forecasting
   - Keywords position 2: acknowledges | promotes | respects | safeguards | by design | upholding
   - Keywords position 3: privacy
   - Exclusions: -no safeguards -lacks safeguards -doesn't promote privacy -does not respect -doesn't respect

3. Privacy
   - Keywords position 1: predictive policing | preventive policing | crime forecasting
   - Keywords position 3: privacy

4. Efficiency, Reliability, Accuracy – positive
   - Keywords position 1: predictive policing | preventive policing | crime forecasting
   - Keywords position 2: effectively | effective | reliable | accurately | accurate | fight crime | prevent crime | increased security | support crime prevention
   - Exclusions: -low efficiency -low reliability -not reliable -not efficient -not accurate

5. Efficiency, Reliability, Accuracy – negative
   - Keywords position 1: predictive policing | preventive policing | crime forecasting
   - Keywords position 2: ineffective | inefficient | not reliable | not accurate | inaccurate | unreliable

6. Efficiency, Reliability, Accuracy
   - Keywords position 1: predictive policing | preventive policing | crime forecasting
   - Keywords position 2: efficiency | reliability | accuracy | efficient | reliable | accurate

7. Legitimacy – positive
   - Keywords position 1: predictive policing | preventive policing | crime forecasting
   - Keywords position 2: legitimacy | legitimate | is legitimate
   - Exclusions: -not legitimate -low legitimacy

8. Legitimacy – negative
   - Keywords position 1: predictive policing | preventive policing | crime forecasting
   - Keywords position 2: not legitimate | illegal

9. Legitimacy

- Keywords position 1: predictive policing | preventive policing | crime forecasting
- Keywords position 2: legitimacy | legitimate | lawfulness | lawful

10. Transparency, Accountability – positive
    - Keywords position 1: predictive policing | preventive policing | crime forecasting
    - Keywords position 2: transparent | accountable | transparency | accountability | oversight
    - Exclusions: -low transparency -no transparency -no accountability -no oversight -lack of oversight -lack of accountability

11. Transparency, Accountability – negative
    - Keywords position 1: predictive policing | preventive policing | crime forecasting
    - Keywords position 2: no | lack | deficient | weak | ineffective | lack of
    - Keywords Position 3: transparent | accountable | transparency | accountability | oversight
    - Exclusions: -Analytica - Autonomous vehicle - CBD Hemp

12. Transparency, Accountability
    - Keywords position 1: predictive policing | preventive policing | crime forecasting
    - Keywords position 2: transparent | accountable | transparency | accountability | oversight

13. Discrimination – positive
    - Keywords position 1: predictive policing | preventive policing | crime forecasting
    - Keywords position 2: no discrimination | impartial | do not discriminate

14. Discrimination – negative
    - Keywords position 1: predictive policing | preventive policing | crime forecasting
    - Keywords position 2: promote discrimination | promote racism | racial bias | disproportionately | inequality
    - Exclusions: -seekersBaiduBenefits

15. Discrimination
    - Keywords position 1: predictive policing | preventive policing | crime forecasting
    - Keywords position 2: discrimination | impartial | bias | discriminate | race | minorities
    - Exclusions: -seekersBaiduBenefits

Cyber operations

1. Privacy – negative
    - Keywords position 1: police | law enforcement | security agencies | enforcement agencies
    - Keywords position 2: cyber operations | cyber technolog*
    - Keywords Position 3: disrupts|impedes|does not acknowledge|not sensitive|harm*|abuse|restrict
    - Keywords Position 4: privacy
    - Exclusions: - foreign policy – war – military - international relations

2. Privacy – positive
    - Keywords position 1: biometric identification | facial recognition | biometrics | biometric* authentication | fingerprint mapping | retina scan* | iris scan*
    - Keywords position 2: cyber operations | cyber technolog*
    - Keywords position 3: acknowledges | promotes | respect | safeguards | by design | upholding
    - Keywords position 4: privacy
    - Exclusions: - foreign policy – war – military - international relations -no safeguards - lacks safeguards -doesn't promote privacy -does not respect -doesn't respect

3. Privacy
    - Keywords position 1: police | law enforcement | security agencies | enforcement agencies
    - Keywords position 2: cyber operations | cyber technolog*
    - Keywords Position 3: privacy
    - Exclusions: - foreign policy – war – military - international relations

4. Efficiency, Reliability, Accuracy – positive
    - Keywords position 1: police | law enforcement | security agencies | enforcement agencies
    - Keywords position 2: cyber operations | cyber technolog*
    - Keywords Position 3: effectively | effective | reliable | accurately | accurate | fight crime | prevent crime | increased security | support crime prevention
    - Exclusions: - foreign policy – war – military - international relations -low efficiency - low reliability -not reliable -not efficient -not accurate

5. Efficiency, Reliability, Accuracy – negative
    - Keywords position 1: police | law enforcement | security agencies | enforcement agencies
    - Keywords position 2: cyber operations | cyber technolog*
    - Keywords Position 3: ineffective | inefficient | not reliable | not accurate | inaccurate | unreliable
    - Exclusions: - foreign policy – war – military - international relations

6. Efficiency, Reliability, Accuracy
    - Keywords position 1: police | law enforcement | security agencies | enforcement agencies
    - Keywords position 2: cyber operations | cyber technolog*

- Keywords Position 3: efficiency | reliability | accuracy | efficient | reliable | accurate
- Exclusions: - foreign policy – war – military - international relations

7. Legitimacy – positive
   - Keywords position 1: police | law enforcement | security agencies | enforcement agencies
   - Keywords position 2: cyber operations | cyber technolog*
   - Keywords Position 3: legitimacy | legitimate
   - Exclusions: - foreign policy – war – military - international relations – not legitimate – low legitimacy

8. Legitimacy – negative
   - Keywords position 1: police | law enforcement | security agencies | enforcement agencies
   - Keywords position 2: cyber operations | cyber technolog*
   - Keywords Position 3: not legitimate | illegal |illegitimate
   - Exclusions: - foreign policy – war – military - international relations

9. Legitimacy
   - Keywords position 1: police | law enforcement | security agencies | enforcement agencies
   - Keywords position 2: cyber operations | cyber technolog*
   - Keywords Position 3: legitimacy | legitimate | lawfulness | lawful
   - Exclusions: - foreign policy – war – military - international relations

10. Transparency, Accountability – positive
    - Keywords position 1: police | law enforcement | security agencies | enforcement agencies
    - Keywords position 2: cyber operations | cyber technolog*
    - Keywords Position 3: transparent | accountable | transparency | accountability | oversight
    - Exclusions: - foreign policy – war – military - international relations -low transparency -no transparency -no accountability -no oversight -lack of oversight -lack of accountability

11. Transparency, Accountability – negative
    - Keywords position 1: police | law enforcement | security agencies | enforcement agencies
    - Keywords position 2: cyber operations | cyber technolog*
    - Keywords Position 3: no | lack | deficient | weak | ineffective | lack of
    - Keywords Position 4: transparent | accountable | transparency | accountability | oversight
    - Exclusions: - foreign policy – war – military - international relations

12. Transparency, Accountability
    - Keywords position 1: police | law enforcement | security agencies | enforcement agencies
    - Keywords position 2: cyber operations | cyber technolog*
    - Keywords Position 3: transparent | accountable | transparency | accountability | oversight
    - Exclusions: - foreign policy – war – military - international relations

13. Discrimination – positive
    - Keywords position 1: police | law enforcement | security agencies | enforcement agencies
    - Keywords position 2: cyber operations | cyber technolog*
    - Keywords Position 3: no discrimination | impartial | do not discriminate

- Exclusions: - foreign policy – war – military - international relations

14. Discrimination – negative
    - Keywords position 1: police | law enforcement | security agencies | enforcement agencies
    - Keywords position 2: cyber operations | cyber technolog*
    - Keywords Position 3: promote discrimination | promote racism | racial bias | disproportionately | inequality
    - Exclusions: - foreign policy – war – military - international relations

15. Discrimination
    - Keywords position 1: police | law enforcement | security agencies | enforcement agencies
    - Keywords position 2: cyber operations | cyber technolog*
    - Keywords Position 3: discrimination | impartial | bias | discriminate | race | minorities
    - Exclusions: - foreign policy – war – military - international relations

Police Hacking

1. Privacy – negative
   - Keywords position 1: police hacking | law enforcement hacking | security agency hacking | enforcement agency hacking | government hacking
   - Keywords position 2: disrupts|impedes|does not acknowledge|not sensitive|harm*|abuse|restrict
   - Keywords Position 3: privacy
   - Exclusions: - -Big Pharma -China CNI -climate change

2. Privacy – positive
   - Keywords position 1: police hacking | law enforcement hacking | security agency hacking | enforcement agency hacking | government hacking
   - Keywords position 2: acknowledges | promotes | respects | safeguards | by design | upholding
   - Keywords position 3: privacy
   - Exclusions: -no safeguards -lacks safeguards -doesn't promote privacy -does not respect -doesn't respect

3. Privacy
   - Keywords position 1: police hacking | law enforcement hacking | security agency hacking | enforcement agency hacking | government hacking
   - Keywords position 3: privacy
   - Exclusions: -Big Pharma -5GAerospaceAfghanistanAfricaAidAir -climate change

4. Efficiency, Reliability, Accuracy – positive
   - Keywords position 1: police hacking | law enforcement hacking | security agency hacking | enforcement agency hacking | government hacking
   - Keywords position 2: effectively | effective | reliable | accurately | accurate | fight crime | prevent crime | increased security | support crime prevention
   - Exclusions: -low efficiency -low reliability -not reliable -not efficient -not accurate

5. Efficiency, Reliability, Accuracy – negative
   - Keywords position 1: police hacking | law enforcement hacking | security agency hacking | enforcement agency hacking | government hacking
   - Keywords position 2: ineffective | inefficient | not reliable | not accurate | inaccurate | unreliable

6. Efficiency, Reliability, Accuracy
   - Keywords position 1: police hacking | law enforcement hacking | security agency hacking | enforcement agency hacking | government hacking
   - Keywords position 2: efficiency | reliability | accuracy | efficient | reliable | accurate

7. Legitimacy – positive
   - Keywords position 1: police hacking | law enforcement hacking | security agency hacking | enforcement agency hacking | government hacking
   - Keywords position 2: legitimacy | legitimate | is legitimate

- Exclusions: -not legitimate -low legitimacy – climate change

8. Legitimacy – negative
   - Keywords position 1: police hacking | law enforcement hacking | security agency hacking | enforcement agency hacking | government hacking
   - Keywords position 2: not legitimate | illegal

9. Legitimacy
   - Keywords position 1: police hacking | law enforcement hacking | security agency hacking | enforcement agency hacking | government hacking
   - Keywords position 2: legitimacy | legitimate | lawfulness | lawful
   - Exclusions: -We and our nominees -Traci Park

10. Transparency, Accountability – positive
    - Keywords position 1: police hacking | law enforcement hacking | security agency hacking | enforcement agency hacking | government hacking
    - Keywords position 2: transparent | accountable | transparency | accountability | oversight
    - Exclusions: -low transparency -no transparency -no accountability -no oversight -lack of oversight -lack of accountability

11. Transparency, Accountability – negative
    - Keywords position 1: police hacking | law enforcement hacking | security agency hacking | enforcement agency hacking | government hacking
    - Keywords position 2: no | lack | deficient | weak | ineffective | lack of
    - Keywords Position 3: transparent | accountable | transparency | accountability | oversight

12. Transparency, Accountability
    - Keywords position 1: police hacking | law enforcement hacking | security agency hacking | enforcement agency hacking | government hacking
    - Keywords position 2: transparent | accountable | transparency | accountability | oversight

13. Discrimination – positive
    - Keywords position 1: police hacking | law enforcement hacking | security agency hacking | enforcement agency hacking | government hacking
    - Keywords position 2: no discrimination | impartial | do not discriminate

14. Discrimination – negative
    - Keywords position 1: police hacking | law enforcement hacking | security agency hacking | enforcement agency hacking | government hacking
    - Keywords position 2: promote discrimination | promote racism | racial bias | disproportionately | inequality
    - Exclusions: -climate change

15. Discrimination
    - Keywords position 1: police hacking | law enforcement hacking | security agency hacking | enforcement agency hacking | government hacking

- Keywords position 2: discrimination | impartial | bias | discriminate | race | minorities

Decision making in justice systems

1. Privacy – negative
    - Keywords position 1: AI | artificial intelligence | algorithm
    - Keywords position 2: justice system | court system | courts | judiciary
    - Keywords Position 3: disrupts|impedes|does not acknowledge|not sensitive|harm*|abuse|restrict
    - Keywords Position 4: privacy
    - Exclusions: - Kinsey Ancient

2. Privacy – positive
    - Keywords position 1: AI | artificial intelligence | algorithm
    - Keywords position 2: justice system | court system | courts | judiciary
    - Keywords position 3: acknowledge | promote | respect | safeguard | by design | upholding
    - Keywords position 4: privacy
    - Exclusions: -no safeguard -lacks safeguards -doesn't promote privacy -does not respect -doesn't respect
3. Privacy
    - Keywords position 1: AI | artificial intelligence | algorithm
    - Keywords position 2: justice system | court system | courts | judiciary
    - Keywords Position 3: privacy

4. Efficiency, Reliability, Accuracy – positive
    - Keywords position 1: AI | artificial intelligence | algorithm
    - Keywords position 2: justice system | court system | courts | judiciary
    - Keywords Position 3: effectively | efficient | reliable | accurately | accurate | fight crime | prevent crime | increased security | support crime preventionKeywords
    - Exclusions: -low efficiency -low reliability -not reliable -not efficient -not accurate

5. Efficiency, Reliability, Accuracy – negative
    - Keywords position 1: AI | artificial intelligence | algorithm
    - Keywords position 2: justice system | court system | courts | judiciary
    - Keywords Position 3: ineffective | inefficient | not reliable | not accurate | inaccurate | unreliable

6. Efficiency, Reliability, Accuracy
    - Keywords position 1: AI | artificial intelligence | algorithm
    - Keywords position 2: justice system | court system | courts | judiciary
    - Keywords Position 3: efficiency | reliability | accuracy | efficient | reliable | accurate
    - Exclusions:

7. Legitimacy – positive
    - Keywords position 1: AI | artificial intelligence | algorithm

- Keywords position 2: justice system | court system | courts | judiciary
- Keywords Position 3: legitimacy | legitimate | is legitimate
- Exclusions: -not legitimate -low legitimacy

8. Legitimacy – negative
    - Keywords position 1: AI | artificial intelligence | algorithm
    - Keywords position 2: justice system | court system | courts | judiciary
    - Keywords Position 3: not legitimate | illegal

9. Legitimacy
    - Keywords position 1: AI | artificial intelligence | algorithm
    - Keywords position 2: justice system | court system | courts | judiciary
    - Keywords Position 3: legitimacy | legitimate | lawfulness | lawful

10. Transparency, Accountability – positive
    - Keywords position 1: AI | artificial intelligence | algorithm
    - Keywords position 2: justice system | court system | courts | judiciary
    - Keywords Position 3: disrupts|impedes|does not acknowledge|not sensitive|harm*|abuse|restrict
    - Exclusions: -low transparency -no transparency -no accountability -no oversight -lack of oversight -lack of accountability

11. Transparency, Accountability – negative
    - Keywords position 1: AI | artificial intelligence | algorithm
    - Keywords position 2: justice system | court system | courts | judiciary
    - Keywords Position 3: no | lack | deficient | weak | ineffective | lack of
    - Keywords Position 4: transparent | accountable | transparency | accountability | oversight

12. Transparency, Accountability
    - Keywords position 1: AI | artificial intelligence | algorithm
    - Keywords position 2: justice system | court system | courts | judiciary
    - Keywords Position 3: transparent | accountable | transparency | accountability | oversight

13. Discrimination – positive
    - Keywords position 1: AI | artificial intelligence | algorithm
    - Keywords position 2: justice system | court system | courts | judiciary
    - Keywords Position 3: no discrimination | impartial | do not discriminate

14. Discrimination – negative
    - Keywords position 1: AI | artificial intelligence | algorithm
    - Keywords position 2: justice system | court system | courts | judiciary
    - Keywords Position 3: promote discrimination | promote racism | racial bias | disproportionately | inequality
    - Exclusions: - Assisted Living

15. Discrimination

- Keywords position 1: AI | artificial intelligence | algorithm
- Keywords position 2: justice system | court system | courts | judiciary
- Keywords Position 3: discrimination | impartial | bias | discriminate | race | minorities