This project that has received funding from the European Union's Horizon 2020 - Research and Innovation Framework Programme, H2020-FCT-2015, under grant agreement no 700381.

Analysis System for GAthered Raw Data

ASGARD

Instrument: Research and Innovation Action proposal

Thematic Priority: FCT-1-2015

D12.5 Societal Impact Analysis

Deliverable number	12.5	
Version:	2.0	
Delivery date:	27.01.2020	
Dissemination level:	PUBLIC	
Classification level:	Non-classified	
Status	FINAL	
Nature:	Report	
Main author(s):	Professor Ruth Fee	Ulster University
Contributor(s):	Professor Martin Haran Dr Esther McGuiness	Ulster University

DOCUMENT CONTROL

Version	Date	Author(s)	Change(s)
0.1	13.10.17	Professor Ruth Fee	
0.2	04.11.17	Professor Ruth Fee	V0.1 reviewed and updated to incorporate feedback from internal and external reviewers.
1.0	30/11/2017		First submission of the deliverable
1.1	09.10.19	Professor Ruth Fee and Dr Lucy Royal-Dawson	Additional literature reviewed and outcomes of workshop incorporated into V0.2
1.2	01.12.2019	Romaios Bratskas (ADI)	Quality review of the deliverable
2.0	03.12.2019	Professor Ruth Fee	Final submission of the deliverable

DISCLAIMER

Contents

Annexes

Figures

1. Introduction

1.1 Overview

Task 12.5 within the Description of Work provides the parameters and framework for D12.5:

> *Task 12.5 Societal Impact Analysis (UU, DCU) [M6-M41]*
>
> *The purpose of this task will be consideration of the societal impact and acceptability of the introduction of the proposed toolset and development of recommendations for mitigating perceived negative consequences. Considered here too will be the implications of deployment of the ASGARD toolset by countries that do not have the same treaty obligations to civil liberties and human rights as EU member states*.

The aim of this document is to provide a context, and set out a proposed methodology, for analysing the societal impact of the ASGARD project. D12.2 set out the privacy constraints of the ASGARD project and the ethical issues of the project and these will not be revisited as part of this paper.

The aim of the ASGARD project is to "provide LEAs with Technological Autonomy by creating a long lasting community of LEAs and the research and development industry, focused on a set of tools and techniques, that facilitate effective collaboration in order to define, develop, share, and evolve open source Big Data technology solutions that will help LEAs prevent and fight against crime and terrorism."

A key challenge of the ASGARD project is to grasp the ethical and legal implications and impacts of the proposed toolset in a potentially sensitive context. Governments and agencies are accumulating data on society as a whole: should algorithms be applied to tease out insights, particularly in the name of preventing crime and terrorism? In the administrative proposal forms for ASGARD, nine out of ten sub-questions in the Societal Impact Table were answered in the positive, including 'yes' to ASGARD research both benefiting but also potentially having a negative impact on society; the reasoning for task 12.5 is detailed in Annex I.

To ensure that the use of the ASGARD toolset is used appropriately (and to understand the meaning of 'appropriate' in the context of LEAs), we need to ask questions such as:

- What do LEAs want from such data?
- What biases are inherent in the algorithms that produce results?
- What legal frameworks should be imposed for positive, just outcomes?

Evaluation, review, oversight, accountability, legal frameworks all seem appropriate if, for instance, the use of Big Data and analytics for use by law enforcement agencies has undesirable impacts on some communities. It is also an important instance, where analytics are used as the basis for calculating "risk assessment scores" for criminal defendants that can be used to make decisions about bail, parole, and

sentencing. [1] The solution cannot be that those responsible for national security, law enforcement, and criminal justice ignore tools that may offer useful insights. The problem is not the concept of data analytics but how it is developed, used, understood and evaluated. We need to make sure that tools are rigorously evaluated against metrics that test not only accuracy and effectiveness but also the disparity of impact and moral questions.

This is a draft deliverable, and the full paper will be completed by month 41 of the project.

1.2 Relation to other deliverables

Deliverable 12.5 is part of the overall Work Package (WP) 12: Social, Ethical, Legal and Privacy (SELP). There are no direct inputs to other deliverables.

1.3 Structure of the deliverable

The document in is five main sections:

- Section 2: An outline of previous studies of the risks and opportunities of Big Data and the potential societal benefits it can bring in the confines of the fight against crime and terrorism
- Section 3: Review of the literature on the challenges of Big Data in the context of Law Enforcement Agencies (LEAs)
- Section 4: Review of the literature on the use of algorithms in forensic analysis by law enforcement agencies
- Section 5: Review of the literature on public confidence, analytics and Big Data
- Section 6: Output from an expert discussions on the societal impact of the introduction of ASGARD-type tools into the work of LEAs.

[1] For reference, see ISO 31000:2009, Risk Management - Principles and Guidelines, International Organisation for Standardization (ISO); ISO/IEC 29134 (project), Information Technology – Security Techniques – Privacy Impact Assessment – Guidelines, International Organisation for Standardization (ISO).

2. Previous studies: the BYTE Project

Previous studies have been carried out on the risks and opportunities of Big Data from a legal perspective, and the potential social benefits it can bring. The Big Data roadmap for cross-disciplinarY community for addressing societal Externalities (BYTE) project, was funded by the European Commission to focus on the positive and negative impacts of Big Data in Europe. As part of the project, a set of case studies was conducted in six different sectors, and roadmaps were developed to provide the necessary research and policy steps to tackle such impacts and develop a socially responsible Big Data economy in Europe.[2]

It is recognized that Big Data has the potential to transform development and accelerate social progress around the world, but there are issues surrounding understanding, ownership, privacy, capacity, measurement and so forth that need further dialogue and discussion.

Cuquet et al (2016)[3] used as one of their case studies "crisis informatics". The results in figure 1 showed positive societal impact around improved awareness and decision making, but more negative impacts around equality, discrimination and privacy. Results of the case studies around privacy and data protection recommended the broadening of Privacy by Design legal/organisational safeguards; collective mechanisms for data protection; and strengthening the role of data protection authorities. In order to address concerns around anti-discrimination practices highlighted by the case studies, the authors recommended a promotion of an anti-discrimination by design approach; a transparent and accountable framework; enhanced co-operation and co-ordination between data protection authorities and equality bodies; and engagement in a societal and political debate on what they termed the "new discriminators" of Big Data.

[2] Cuquet, M., Vega-Gorgojob, G., Lammerantc, H., Finnd, R. and Hassane, U. (2016) *Societal impacts of Big Data: challenges and opportunities in Europe* https://arxiv.org/ftp/arxiv/papers/1704/1704.03361.pdf
[3] *Op cit*

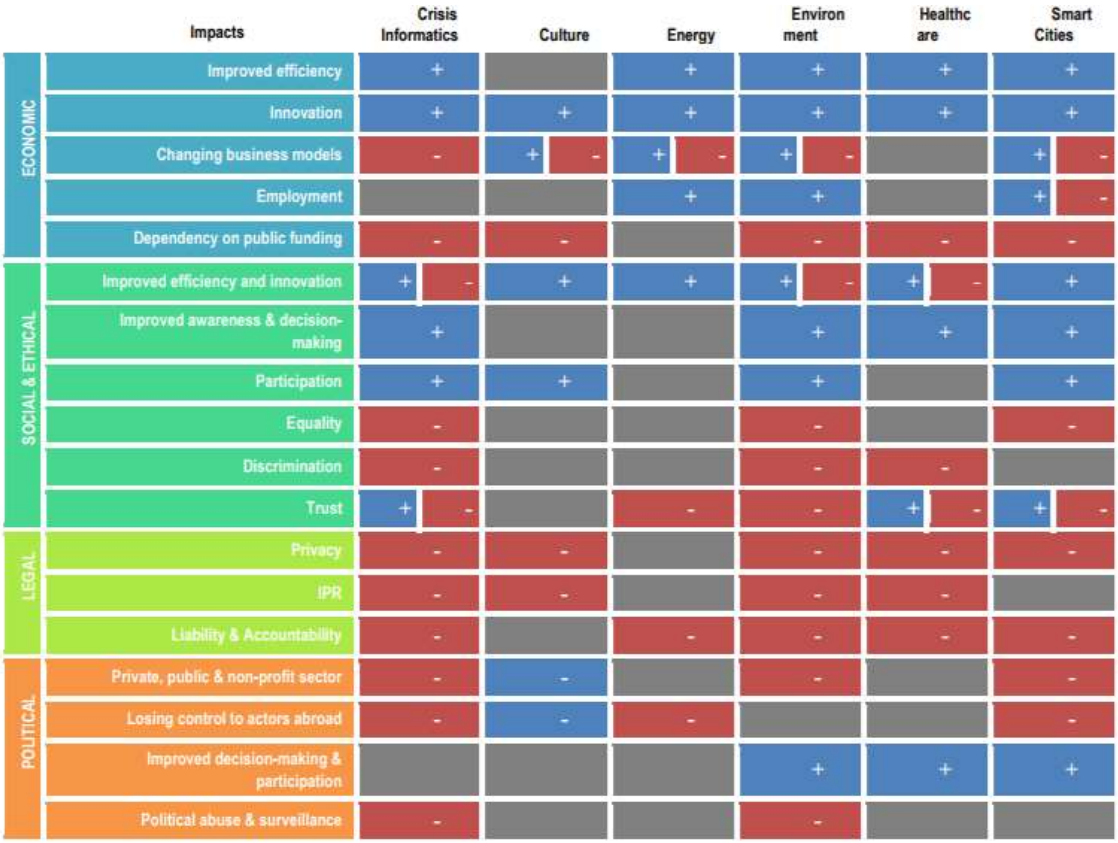

Figure 1: Societal impacts grouped as economic, social and ethical, legal, and political areas, and their positive or negative presence in the BYTE case studies[4]

[4] *Op cit*

2.1 Social benefits of Big Data – the BYTE Project

The use of large data sets for data analytics, predictive analytics and deep learning does not only pose legal challenges and opportunities, but also carries significant potential societal benefits when data is used responsibly. The BYTE project identified six areas where the use of Big Data can result in societal benefit, with the impact on BYTE sectors depicted in figure 2.

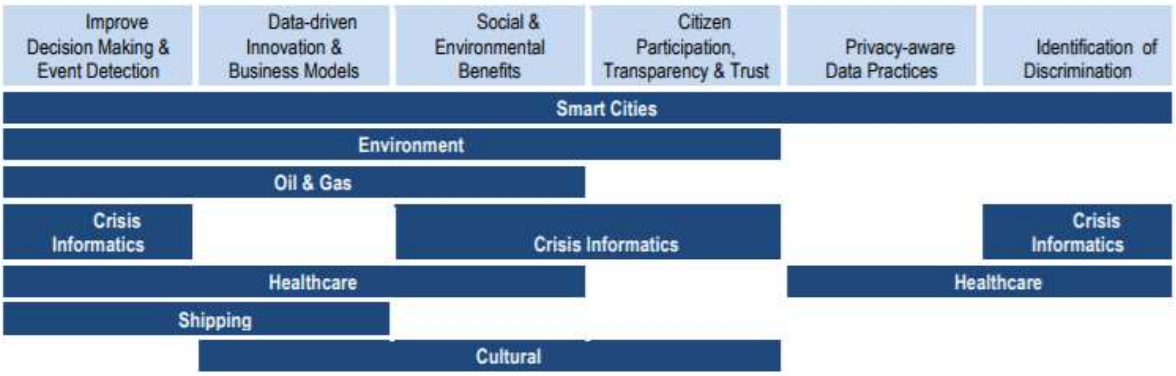

Figure 2. Mapping of the social benefits of Big Data against different sectors[5]

These six areas are improved decision making and event detection, including efficient resource allocation; data-driven innovations, including new business models; direct social, environmental and other citizen benefits; citizen participation, transparency and public trust; privacy-aware data practices; and Big Data for identifying discrimination.

2.1.1 Improved decision making and event detection

For the ASGARD project, improved decision making and event detection, including efficient resource allocation is of particular relevance. Stakeholders from across the BYTE case studies noted that one of the key societal benefits of big data analytics was improvement in decision-making and situational awareness. This also includes improvements in efficiency of resource allocation based on a better understanding of the situation and an ability to target resources. This improved decision-making and event detection, was evident in almost all, of the BYTE case studies.

[5] *Op cit*

2.1.2 Social and ethical benefits for citizens

Innovative practices surrounding Big Data may also result in social and ethical benefits for citizens and the environment, either through improved services, better situational awareness, or personalised or targeted services. These benefits ranged from environmental protection to improved life chances to personalised government services for individual citizens. These benefits were seen across the crisis, environmental, oil and gas, health and smart cities case studies in the BYTE project, but for ASGARD, the benefits could be an increase in the levels of security through enhanced prevention, and fighting of crime and terrorism.

2.1.3 Privacy aware data practices

Despite the fact that privacy is often constructed as a challenge to innovation (Michael and Miller, 2013), the BYTE research found that in some case studies one of the unexpected societal benefits of Big Data practice, has been increased attention to privacy and data protection issues by practitioners. For example, practitioners in crisis informatics, environmental science and healthcare have been particularly sensitive to potential privacy issues and have constructed good practice strategies to both protect privacy and provide innovative, data dependent services. For the ASGARD project, these aspects are complex, although it is clear from studies identified in section 3 below that LEAs are aware of privacy concerns but there is a lack of guidance on the application of the legal framework. The benefits of Big Data will be meaningless if legal frameworks are not complied with, resulting in evidence that is not admissible and could ultimately compromise successful prosecutions.

2.1.4 Big Data for identifying discrimination

Finally, despite the serious and significant potential for Big Data to result in discrimination, the BYTE project has also found that in certain circumstances, Big Data can be used to identify and, consequently, combat discriminatory practices. Systems like DCUBE[6] can be used to identify discriminatory classification rules from the historical data in order to intervene in these practices. Other pre-and post-processing techniques, can also be used to remove or compensate for discrimination within training datasets, including massaging the data, reweighting particular variables, resampling or applying model correction methods. Each of these methods can both combat discrimination as well as ensure responsible and ethical data practices moving forward within the Big Data landscape. These methods will be tested as part of the development of the ASGARD toolset.

In summary, the BYTE project showed through the case studies that the positive social impact of improved, evidence-based decision making is present in several case studies. In order to capture these benefits, several best practices have been suggested by the BYTE project. To address

[6] Ruggieri, S. & Pedreschi, D. & Turini, F. (2010) *DCUBE: Discrimination discovery in databases*. Proceedings of the ACM SIGMOD International Conference on Management of Data. 1127-1130. 10.1145/1807167.1807298.

discrimination, equality and trust, privacy-by-design methods should be extended to anti-discrimination-by-design and analogous approaches, and transparency and new accountability frameworks need to be based both on legislation and on a data protection framework.

Overall, policy makers, regulators and stakeholders all have an important role in updating legal frameworks, promoting Big Data practices, and developing and incorporating tools into the Big Data design and practice that address societal concerns.

3. Law Enforcement Agencies (LEAs) and the challenges of Big Data

Annex 1 of this paper sets out the rationale of conducting an analysis of the societal impact of the use of the ASGARD toolset by law enforcement agencies. There is limited research available in English that explores the use of Big Data analytics for policing and such research appears to be primarily in a UK context, although the findings can be transferred and applied to other jurisdictions. Babuta (2017)[7] carried out interviews with 25 serving UK police officers and staff, as well as experts from the technology sector and academia, that provided new insights into the limitations of the police's current use of data and the police's priorities for expanding these capabilities.

Babuta identified four key priorities in which Big Data technology could be applied to policing (specifically in a UK context). First, predictive crime mapping could be used to identify areas where crime is most likely to occur, allowing limited resources to be targeted most efficiently. Second, predictive analytics could also be used to identify the risks associated with particular individuals. This includes identifying individuals who are at increased risk of reoffending, as well as those at risk of going missing or becoming the victims of crime. Third, advanced analytics could enable the police to harness the full potential of data collected through visual surveillance, such as CCTV images and automatic number plate recognition (ANPR) data. Fourth, Big Data technology could be applied to open-source data, such as that collected from social media, to gain a richer understanding of specific crime problems, which would ultimately inform the development of preventive policing strategies.

Further to the considered application potential of Big Data analytics within UK policing it is noteworthy that Babuta's research also served to identify a number of practical and organisational barriers to implementing these technologies. Most significantly, the lack of coordinated development of technology across UK policing is highly problematic for Big Data, which relies on effective nationwide data sharing and collaboration. Financial cuts in recent years have also severely hindered technological development, as the majority of police IT budgets is spent supporting existing legacy systems, with

[7] Babuta, A. (2017) *Big Data and Policing An Assessment of Law Enforcement Requirements, Expectations and Priorities* RUSI Occasional Paper. https://rusi.org/sites/default/files/rusi-bigdata-press-2017.pdf

little funding available to invest in new technology. The ASGARD project is therefore a critical development within this context.

The author also referred to the significant legal and ethical constraints governing the police's use of data. It was noted that the complex ethical questions and the ethical implications of Big Data, remain poorly understood by law enforcement agencies. As the use of such technology becomes increasingly widespread, legislative frameworks must expand to incorporate 'new rules to regulate the societal cost of our new tools without sacrificing their undeniable benefits'.

Ethical concerns around the use of data typically focus on the collection, analysis and dissemination of personally identifiable information. The EU General Data Protection Regulation introduced further rules governing the collection and use of personal data, and Directive EU 2016/68018 legislates for the processing of personal data for policing purposes. Both came into effect in May 2018, and both apply to the automated analysis of personal data as well as manual analysis. Crucially, data protection legislation in all its forms, does not apply to anonymised datasets.[8] For this reason, the Law Enforcement Directive suggests that organisations should aim to 'pseudonymise' datasets as early as possible, to facilitate 'the free flow of personal data within the area of freedom, security and justice'. Pseudonymisation is defined as: the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person.

Data protection legislation restricts large-scale exploratory analysis of personal data, such as the predictive risk assessment methods discussed in Chapter III. According to Babuta (2017), pseudonymisation is likely to be the only way to perform Big Data analytics on personal datasets while complying with data protection laws. For information gleaned from open sources on the internet, the author notes that "privacy is not a binary concept, but rather 'virtually all information exists in intermediate states between completely public and completely private'";[9] when such analysis is justified and necessary for a specific policing purpose, there is no reason why investigators should not be able to carry it out themselves. Data protection legislation is highly complex and nuanced, and at present law enforcement agencies practitioners have limited sources of accessible and practical guidance on what constitutes the appropriate use of data.

3.1 Bulk data collection and retention – and data manipulation

An earlier publication by RUSI reports on the Independent Surveillance Review conducted at the request of the British deputy prime minister after the allegations of mass surveillance conducted by

[8] *Op cit*

[9] *Op cit* citing Richards and King (2014) 'Big Data Ethics,' p413.

UK and USA governments of their people. [10] While the focus of much public concern relates to bulk data collection, the panel behind the RUSI (2015) report believes it is important to distinguish between the relative impact on privacy on each of the processes of data collection, retention and analysis. Privacy issues need to be considered afresh at each stage.

The panel was persuaded by the argument that intelligence agencies in particular, will always need to conduct both targeted (that is, specifying the individuals or premises to be covered by the warrant) and untargeted data collection (recognising that even untargeted collection must be specifically aimed at achieving an authorised mission or intelligence requirement). Targeted data collection will be needed when the agencies have identified a subject or subjects of interest and require further information on them, if only to confirm whether or not they pose a threat. Some degree of untargeted data collection, involving the collection of data in bulk, may sometimes be required, especially given the nature of modern communications. This is particularly relevant to the ASGARD project, using a range of bulk open source data.

The RUSI (2015) report accepts that some critics will remain convinced that untargeted data collection as a principle is unacceptable, but the ability of the LEAs to collect data in bulk may in some instances be necessary when there is no viable alternative for them to identify potential and unknown threats, particularly online. However, the Snowden disclosures show how such data collection can be undertaken without public awareness or consent. Such awareness and consent are crucial, as are robust oversight mechanisms to reassure the public that capabilities are not being misused or abused. There are further concerns about what happens to an individual's data after it has been collected, in particular the circumstances in which this data is interrogated and analysed and for how long data is kept. Policies on data retention must be subject to regular review by oversight bodies to ensure they remain proportionate (and, as noted above, oversight mechanisms must have the technical knowledge to monitor this effectively). [11]

[10] Royal United Services Institute for Defence and Security Studies (2015) *A Democratic Licence to Operate Report of the Independent Surveillance Review*, Whitehall Report 2-15, July 2015. https://rusi.org/sites/default/files/20150714_whr_2-15_a_democratic_licence_to_operate.pdf

[11] The specific issue of re-evaluating the necessity of how long data is kept was taken up in the Europol annual inspection report 2012 which found: "*The processing of personal data of a group of 96 non-violent anarchists reported by the contributor as not representing any danger is not - and has never been - in compliance with the opening order." …." Retaining data over 5 yrs without any check on their relevance, and reasoning the further processing as part of dealing with a backlog of data to be processed, does not constitute a proper evaluation of the necessity of the retention of these data".* Therefore, building in data retention periods at the start of a proposed measure and re-evaluating it periodically is essential to ensure compliance with a persons' right to a private life and data protection law.

3.2 Collaboration between the public and private sectors

The private sector is highly internationalised and evolves rapidly, yet its role tends to be overlooked in debates over privacy and security. Given that commercial organisations are the largest generators and guardians of citizens' data, it is important to understand the types and volumes of data collected and what is subsequently done with it. The collection and manipulation of bulk data is not something unique to government, but rather a pervasive technique which a growing number of organisations, both in the private as well as the public sector, now used to interact with the public as citizens and customers (RUSI, 2015).

4. Law enforcement agencies and the use of algorithms in forensic analysis

Some of the ASGARD tools use algorithmic machine learning. An algorithm is 'a sequence of instructions that are carried out to transform the input to the output' (Alpaydin, 2016).[12] This computer science employs coding that learns and adapts itself for the given task using the input data. The algorithm develops according to the input data. The algorithm is thus highly dependent on the input data.

Three main purposes of algorithmic analysis within the police context have been identified:

> "(i) predictive policing on a macro level incorporating strategic planning, prioritisation and forecasting;
>
> (ii) operational intelligence linking and evaluation which may include, for instance, crime reduction activities; and
>
> (iii) decision-making or risk-assessments relating to individuals." (Oswald and Grace, 2016)[13].

The ASGARD toolset is designed for the third purpose, but it is possible the tools may be used for the other two purposes too. Ethical concerns arise from the use of algorithms in policing related to poor decisions, lack of transparency, bias, discrimination and an impact on the decision-making process itself. Mittelstadt *et al.* (2016)[14] categorised the concerns as follows:

> '(1) inconclusive evidence leading to unjustified actions;

[12] Alpaydin, E. (2016) *Machine Learning.* MIT Press, Mass.

[13] Oswald, M. and Grace, J. (2016) 'Intelligence, policing and the use of algorithmic analysis: a freedom of information-based study,' *Journal of Information Rights, Policy and Practice*, 1(1).

[14] Mittelstadt, B.D., Allo, P., Taddeo, M., Wachter, S. and Floridi, L. (2016) 'The ethics of algorithms: Mapping the debate,' *Big Data & Society*, July – December, pp1 – 21.

(2) inscrutable evidence leading to opacity;

(3) misguided evidence leading to bias;

(4) unfair outcomes leading to discrimination; and

(5) transformative effects leading to challenges for autonomy and informational privacy.'

Not all of these ethical concerns are related only to algorithmic analysis and they may well arise in relation to other decision-making processes, such as human decision-making. Here we are interested in the implications of algorithmic analysis on law enforcement using the ASGARD tools. It may be a worthwhile exercise to consider the likelihood of outcomes produced by the tools that may be troubled by these ethical concerns.

4.1 Bias from input data and 'bias by design'

Of immediate concern is the possibility of biased outcomes arising from misguided evidence. O'Neil (2016)[15] demonstrates how the use of selected existing data from policing databases in crime hot-spot prediction machine learning software in Philadelphia skewed the analysis, resulting in a pernicious feedback loop. Some neighbourhoods were predicted to be more likely to be crime hot-spots, which focused police resources on them, which resulted in more arrests and so more policing, while overlooking other neighbourhoods. Oswald *et al.* (2018)[16] point out that the use of historic data to train an algorithm makes the assumption that the decisions of the past were all good or that the same decisions would be made now. They suggest that bias inherent in the data should be sought out and controlled before it is used to train the algorithm and not doing so risks embedding bias deep with the model, obscured from accountability. The inclusion of irrelevant data too may result in bias, such as the inclusion of data on individuals' spent convictions when they have no relevance to the investigation. Biased results may also arise from insufficient or incomplete data, such as when a minority population is less fully represented in a population; or the use of in appropriate data attributes, such as benefit status, or the use of proxies, such as post-codes (HCSTC, 2018).[17]

Another concern is the implicit, unwitting use of value-laden judgements within the algorithm. For example, the coder's decision to permit a false negative result (not resulting in the correct target) but

[15] O'Neil, C. (2016) *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Penguin Books. p86.

[16] Oswald, M., Grace, J., Urwin, S. and Barnes, G.C. (2018) 'Algorithmic risk assessment policing models: lessons from the Durham HART model and 'experimental' proportionality,' *Information & Communications Technology Law,* 27(2), pp223-250. p235

[17] House of Commons Science and Technology Committee [HCSTC] (2018) *Algorithms in decision-making. Fourth Report of* Session 2017-19. Available at: https://www.parliament.uk/business/committees/committees-a-z/commons-select/science-and-technology-committee/inquiries/parliament-2017/algorithms-in-decision-making-17-19/. Para 37-38.

reject a false positive result (resulting in the incorrect target) may have consequences over time on how the algorithm performs (Oswald *et al.*, 2018).[18] Some commentators have noted that the lack of diversity amongst the writers of algorithms may also unwittingly introduce bias into way algorithms operate (HCSTC, 2018).[19]

Babuta and Oswald (2019) report on an in-depth consultation they conducted for the UK government's Centre for Data Ethics and Innovation with the aim of making recommendations for a code of practice for the policing sector on guarding against bias in algorithmic and other forms of data analytics in 2020.[20] In their Briefing Paper '*Data analytics and algorithmic bias in policing*,' they highlight the risk of bias arising from algorithmic analysis leading to unfair discrimination of protected characteristics and also from the organisational, operational or legal context. They also stress the need to update the legal framework for the use of analytics in step with policy and regulation developments. Measures for scrutiny, regulation and enforcement, they report, are needed to accompany these developments.

Babuta and Oswald (2019) draw attention to the link between a tool's effectiveness and accuracy and its legality and ethics. They note the importance of evaluating a tool's accuracy and effectiveness so that its continued use can be assessed for its legality and justified use in a policing function. With the haphazard and uncoordinated introduction of analytical or algorithmic tools in policing, it is essential to determine the effectiveness and accuracy of systems and develop 'a clearer legal, policy and regulatory framework to ensure proportionate and ethical use of this increasingly powerful technology' (p8). They also highlight how bias can be introduced at the different stages of the decision to deploy an analytical tool: at problem identification stage when data-driven assessments may inaccurately target a sub-population, for example; at design and testing stage when incomplete or biased training data are used as input data or proxies are used; at deployment when some data over others are selected leading to a skewed analysis or when automation bias results in a tendency to over-rely on outputs from an automated analysis.

Thus, bias in algorithms can arise not only from unrepresentative or biased data used to train the algorithm, but also from incomplete or incorrect data and from the design or construction of the algorithm.

[18] Oswald *et al.* at note 14. p236
[19] As at 14. Para 43
[20] Babuta, A. and Oswald, M. (2019) *Data analytics and algorithmic bias in policing,* Royal United Services Institute for Defence and Security Studies. Available at: https://rusi.org/publication/briefing-papers/data-analytics-and-algorithmic-bias-policing.

4.2 Regulation, oversight and accountability

The Law Society of England and Wales make recommendations in their 2019 report for clarity and explicit explanations of the lawful basis for algorithmic systems in the criminal justice system.[21] These should include an explicit statement of the lawful basis of the use of facial recognition capabilities. To increase the protection of biometric data, they recommend the role of the Biometrics Commissioner be bolstered with additional powers to scrutinise and additional resources. Spanning all uses of algorithmic analyses in the criminal justice system, they recommend there should be a range of mechanisms of improved oversight that may include frequent reviews, expanded capacity of the Information Commissioners to examine algorithms, a code of practice, complaints mechanisms and national register of algorithmic systems. Alongside these governance and accountability measures, they recommend strengthening data protection, privacy and other human rights when algorithms are used in the justice sector. The use of data protection impact assessments, guidance on user logging systems and safeguards for the prominence of human intervention over automated decisions are three such suggestions. With regards to the fairness and transparency for the use of algorithms in the justice sector, the Law Society recommends that equality impact assessments, including socio-economic equality, are conducted as a pre-deployment requirement. At the development stage too, they strongly recommend that the specifications of system, such as the problem definition, are never outsourced and are developed to 'allow for maximal control, amendment and public-facing transparency, and be tested and monitored for relevant human rights considerations' (p7). To achieve this, they suggest 'human rights by design,' the development of explanation capacity to help with the assessment of whether a decision is justified and facilitate better understanding and scrutiny of the systems.

4.2.1 Legislation

The evidence put forward as part of the RUSI (2015) review indicates that the public in Britain is generally supportive of the work and expertise of the LEAs and of the requirements of intelligence-led policing. There is a problem of trust in the system of oversight, and particularly the lack of popular visibility of the oversight arrangements that currently exist. The report outlined ten enduring tests that government, and the public, should apply when considering all future legislation relating to the conditions under which the police and intelligence and security agencies can intrude upon the privacy of the citizen.

[21] The Law Society of England and Wales (2019) *Algorithms in the criminal justice system – A report by the Law Society Commission on the Use of Algorithms in the Justice System.* Available at: https://www.lawsociety.org.uk/support-services/research-trends/algorithm-use-in-the-criminal-justice-system-report/

1. Rule of law: All intrusion into privacy must be in accordance with law through processes that can be meaningfully assessed against clear and open legislation, and only for purposes laid down by law. [22]

2. Necessity: All intrusion must be justified as necessary in relation to explicit tasks and missions assigned to government agencies in accordance with their duly democratic processes, and there should be no other practicable means of achieving the objective.[23]

3. Proportionality: Intrusion must be judged as proportionate to the advantages gained, not just in cost or resource terms but also through a judgement that the degree of intrusion is matched by the seriousness of the harm to be prevented.[24]

[22] In the case of *MM v United Kingdom Appl. No. 24029/07 (ECtHR 13 November 2012),* the ECtHR set out the criteria that must be met for an act or activity to be 'in accordance with the law'. An activity must:

- have some basis in domestic law and be compatible with the rule of law; and
- the law must be adequately accessible and foreseeable, that is,
- formulated with sufficient precision to enable the individual to regulate his or her conduct

[23] Art. 8(1) of the ECHR provides that:
'Everyone shall have the right to respect for his private and family life, his home and his correspondence.'
However, the right is not absolute and Art. 8(2) sets out the grounds the State may interfere with an individual's right to privacy:
'There shall be no interference by a public authority with the existence of this right except such as is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or the economic well-being of the country, for the prevention or detection of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others.' The ECtHR has set out three criteria which must be satisfied to ensure that any interference is in compliance with Art. 8(2). So an interference must be:

- in accordance with the law,
- in pursuit of one of the legitimate aims set out in Art. 8(2), and
- necessary in a democratic society

There is a wealth of case law in this area which could assist LEAs in developing policy guidelines. See for example; *S & Marper v United Kingdom* Appl. Nos. 30562/04 and 30566/04 (ECtHR 4 December 2008) Par. 101; *Khelili v Switzerland* Appl. No. 16188/07 (ECtHR 18 October 2011*); Klass and others v Germany* Appl. No. 5029/71 (6 September 1978); *Leander v Sweden* Appl. No. 9248/81 (ECtHR 26 March 1987); *Huvig v France* Appl. No. 11105/84 (ECtHR 24 April 1990); *Z v Finland* Appl. No 22009/93 (ECtHR 25 February 1997*); K & T v Finland* Appl. No. 25702/94 (12 July 2001)

[24] Two notable cases heard by the ECtHR involving the issue of proportionality in the area of privacy law are *Z v Finland,* Appl. No. 22009/93 (ECtHR 25 February 1997); and *S & Marper v United Kingdom, op cit,* In the case of *S & Marper*, the applicants complained that the retention of their DNA and fingerprint samples by the Police constituted an unjustified interference with their Art. 8 rights. In *Z*, the issue was that the applicant's personal information (including her health status) was publicly disclosed.

4. Restraint: It should never become routine for the state to intrude into the lives of its citizens. It must be reluctant to do so, restrained in the powers it chooses to use, and properly authorised when it deems it necessary to intrude. [25]

5. Effective oversight: An effective regime must be in place. Effectiveness should be judged by the capabilities of the regime to supervise and investigate governmental intrusion, the power it has to bring officials and ministers to account, and the transparency it embodies so the public can be confident it is working properly. There should also be means independently to investigate complaints.

6. Recognition of necessary secrecy: The 'secret parts of the state' must be acknowledged as necessary to the functioning and protection of the open society. It cannot be more than minimally transparent, but it must be fully democratically accountable.

7. Minimal secrecy: The 'secret parts of the state' must draw and observe clear boundaries between that which must remain secret (such as intelligence sources or the identity of their employees) and all other aspects of their work which should be openly acknowledged. Necessary secrecy, however, must not be a justification for a wider culture of secrecy on security and intelligence matters.

8. Transparency: How the law applies to the citizen must be evident if the rule of law is to be upheld. Anything that does not need to be secret should be transparent to the public; not just comprehensible to dedicated specialists but clearly stated in ways that any interested citizen understands.

9. Legislative clarity: Relevant legislation is not likely to be simple but it must be clearly explained in Codes of Practice that have Parliamentary approval, are kept up-to-date and are accessible to citizens, the private sector, foreign governments and practitioners alike.

10. Multilateral collaboration: Government policy on intrusion should be capable of being harmonised with that of like-minded open and democratic governments.

These "enduring tests" are all relevant in the context of the ASGARD project and how they should be applied to the development of the toolset needs to be considered.

[25] A LEA must have a legal framework (codified or common /statue law) to ensure that the powers it exercises are legitimate. It is of particular importance in a codified jurisdiction that the LEA has a legal basis to exercise specific powers to carry out its functions in pursuit of that legitimate aim. For example, in Dudgeon v United Kingdom, Appl. No. 7525/76 (ECtHR 23 September 1981), it was not disputed that the police acted in accordance with the law, or that they were pursuing a legitimate aim. However, the police had failed to demonstrate that the steps they took to intrude on Mr Dudgeon's private life were 'necessary in a democratic society'.

4.2.2 Accountability

The capacity for challenging an outcome that is a direct result of an algorithmic process needs to be developed in line with the accountability structures for challenging outcomes derived from analogue methods. Determining who is responsible for any decision taken is a key factor in this process. HCSCT (2018) stresses the need for robust accountability and transparency structures to permit scrutiny and challenges to decisions. HCSTC recognise it is difficult to challenge algorithmic results, and to cross-examine results in the way normal evidence can be.

4.2.3 Assigning responsibility

At the development stage of an algorithm, the developer may not know the future intended use of the algorithm and user may not know how the algorithm operates. The locus of accountability is not clear (HCSTC, 2018).[26] Since algorithms and automated decisions are not accountable to anyone, mechanisms for challenging algorithms and channels for accountability need to be considered. The Royal Academy of Engineers recommends that governance and accountability be considered at the development stage to ensure that the correct assumptions about how the algorithm will be used and possibly challenged can be incorporated into its development.[27]

The concerns over the lack of regulation, oversight, accountability measures, transparency, explanations and human rights and equality impact assessments associated with the LEA use of algorithmic systems suggest that the roll-out of the ASGARD tools would benefit from documentation aimed at the public and recommendations for mitigation strategies for these concerns aimed at LEAs.

4.2.4 The challenge of national jurisdictions

If the Internet, by its very nature, straddles all the national legal jurisdictions of its users, the fact remains that law-enforcement and intelligence organisations – as agents of the state – are by definition subject to jurisdictional boundaries. The European Court of Human Rights provides binding instruments that govern certain aspects of the legal frameworks of its signatory members. This, however, does not cover non-European states and only some of the most relevant aspects of interception and surveillance among its member states. Beyond that, a number of reviews have been conducted and guidelines suggested through the UN, the EU and the Council of Europe to suggest harmonisation measures that would bring law, practice and the culture of security closer together

[26] As at 15. para 46.

[27] The Royal Academy of Engineering, 2018, *Written evidence submitted by The Royal Academy of Engineering,* Available at: http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/science-and-technology-committee/algorithms-in-decisionmaking/written/69070.pdf)

between states that are still catching up with the implications of Internet technology on their human rights as well as their security concerns.[28]

4.3 Impact of using algorithms on the autonomy of the end user

Oswald *et al.* (2018)[29] warn that the end users of algorithms may experience a degradation in their autonomy when automated tools are used. The lack of transparency in how outputs are reached pushes the end user to the periphery of understanding how results are arrived at. It becomes difficult for the end user to justify the result and even to challenge it. His or her discretion may become fettered to whatever the computer says. This has consequences for accountability when it is not justifiable to simply point to the 'black box' of the computer.

5. Public confidence, analytics and Big Data

The ASGARD toolset is designed to define, develop, share, and evolve open source Big Data technology solutions that will help LEAs prevent and fight against crime and terrorism. It is noted above the challenges LEAs must overcome in order to utilise Big Data effectively, including the impact on society, ethics and privacy and law. Understanding the levels of public confidence and trust in the use of analytics in general and specifically by LEAs will be important for ensuring their acceptability and perceived validity in investigative work.

5.1 Public opinion

As the market-research organisation Ipsos MORI points out, 'there is no one public opinion on data privacy'.[30] In analysing the results of various polls, studies and surveys conducted over the last three years, Ipsos MORI reported that there is significant variation in public awareness of how data are collected, used and shared; in public understanding of the parameters of the debate; and in how concerned different people are by threats to their personal privacy. These concerns are also specific to each situation – people do not tend to simply make a general 'trade-off' between privacy and security – and opinions can change depending on different data use, data users and data purposes. Research also shows that, while people may be concerned in general terms, data-privacy issues are not at the forefront of their thoughts, and their behaviour may not reflect stated levels of concern. Indeed, Ipsos MORI notes that 'stated concern about data privacy and how people actually behave are barely nodding acquaintances'.

[28] Council of Europe, Commissioner for Human Rights (2015) *Democratic and Effective Oversight of National Security Services*, Issue Paper, Council of Europe.
[29] As at 16. p237.
[30] Ipsos MORI, *Understanding Society: The Power and Perils of Data*, 2014, p. 2.

The RUSI (2015)[31] report argued that there is reason to suspect that the British public are most concerned by data collection and use by the private sector. According to the 2014 UK TRUSTe Privacy Index, 20 per cent of those who said they were concerned by online privacy said that this was caused by reports of government surveillance; 60 per cent were concerned because of businesses sharing personal information with other companies.

5.2 Perceptions of Law Enforcement Agencies

The RUSI (2015) report notes that the public's support for legitimate state law-enforcement and security and intelligence work is crucial, and the police and intelligence agencies themselves are the first to acknowledge that they require public consent – it underpins their licence to operate. Even if it is universally accepted that the agencies must keep some operational details of their work secret, the public must support in principle what the agencies do, and be confident they are acting within a legal framework. The public must also remain confident in the accountability and oversight mechanisms which verify that the agencies are operating within justifiable moral, ethical and legal limits, and their work carried out in the public interest. [32]

This is further supported by a YouGov poll of January 2015 which asked whether the public thought the security services did or did not need more access to the public's communications (such as e-mails and phone calls) in order to effectively fight terrorism. The majority (52 per cent) believed they did need more access, compared to 31 per cent which believed that they already have all the access they need or more than they need, while 17 per cent did not know. Overall, trust in intelligence agencies also appears to be high, even when compared to the police. In the same YouGov poll, 63 per cent of respondents said they would have trust in the intelligence services to behave responsibly with information obtained using surveillance powers, compared to 29 per cent who said they would not have trust. For the police, 50 per cent claimed they would trust the police to behave responsibly, compared to 42 per cent who said they would not have trust.[33]

5.3 The public's awareness of data collection and use

The RUSI (2015) report states that it is reasonable to suggest that the public's perceptions of surveillance, police, the agencies and oversight would change if they were more aware of some of these issues and, in particular, if they were aware of how much of their data is collected and used. It is not clear the extent to which the public fully appreciate the scale of data collection permitted within

[31] Royal United Services Institute for Defence and Security Studies (2015) A Democratic Licence to Operate Report of the Independent Surveillance Review, Whitehall Report 2-15, July 2015. https://rusi.org/sites/default/files/20150714_whr_2-15_a_democratic_licence_to_operate.pdf

[32] For these reasons, consideration must also be given to the legal basis of the action, with particular regard for Art 8 (2) European Convention on Human Rights.

[33] YouGov/Sunday Times, 'Survey Results', 15–16 January 2015, <https://d25d2506sfb94s.cloudfront.net/cumulus_uploads/document/wt26kxdn72/YG-Archive-Pol-Sunday-Timesresults-160115.pdf>

digital society. Much of this data collection occurs without the public even realising it. However, at all times, it is incumbent on LEAs, to ensure that the processing of information used to detect or deter criminal activity, is undertaken lawfully. The report summarises the significant controversy over the interception of citizens' data as revolving around three key questions:

- Are the public aware of who can collect their data and for what purpose?
- Do the public have confidence in the legislative and governance frameworks which govern data usage?
- Are the public satisfied with the authorisation, accountability and oversight checks that are in place?

5.4 Confidence and oversight

Public confidence in the acquisition and retention of data rests on the credibility and practicality of the legal and oversight frameworks that govern it.

The RUSI (2015) report states that on the 'security' side, LEAs argue that, in order to stay one step ahead of increasingly capable adversaries, they must have a particular set of advanced and potentially intrusive capabilities. On the 'privacy' issue, civil-liberties and privacy advocates, believe that the capabilities of the agencies are disproportionate to the threat, and that the oversight mechanisms that hold them to account are inadequate**.** It was within this context that the disclosures by Edward Snowden emerged. The information brought sharp focus to the debate and entrenched positions on both sides. Allegations of bulk data collection raised specific legal questions on the remit and oversight of the LEAs in many countries, including the UK. Open societies have to protect themselves, but the parts of the state entrusted with significant powers must be carefully regulated and held to a high level of accountability.

6. Expert commentary on societal impact

To provide current commentary and reflection on the societal, ethical, privacy and legal aspects of the ASGARD toolset, a workshop was convened with discussants from a range of relevant professional backgrounds. The 24 participants came from human rights, child rights and privacy rights organisations, technology ASGARD partners, LEA partners of ASGARD and from other LEA forces, legal professionals and academics. A full list of participants is given in Annex 2.

The aim of the workshop was to provide commentary and critique on the likely impact on society of a suite of investigative and forensic analysis tools such as those being developed for use by law enforcement agencies by ASGARD.

6.1 Workshop structure

To maximise transport options and representation from the various professional backgrounds in Europe, UK and the island of Ireland, the workshop was held at Dublin City University over one full day.

To provide the background to ASGARD and the tools, presentations were delivered as follows:

1. Introduction to ASGARD – Bernardo Pacheco, INOV
2. Demonstration of the tools with a use-case – Owen Corrigan, DCU, and Stephan Veigl, AIT
3. Law enforcement view of how the tools will be used – Armin Vogl, BMI and Fredrik Johansson, FOI
4. Legal, privacy and ethical issues in developing the tools and privacy by design – Ruth Fee, Ulster University

The participants worked in groups to discuss the privacy, data protection and ethical concerns about the development, introduction and use of AI and machine-learning analytical tools intended for use in criminal investigations, criminal behaviour forecasting and surveillance. They interrogated the use of such tools by law enforcement agencies, and what needs to be in place to safeguard against things going wrong.

6.2 Commentary from participants

The output of the discussion parallels many of the concerns discussed in the literature in the previous sections, reflecting the current trend in the concerns around the use of AI and machine-learning analytical tools in the police context. There were, in addition, some concerns and different emphases placed on concerns that are less well reported in the literature. The next two sections summarise the similarities and then the divergencies. A summary of the entire discussion is contained in Annex 2.

6.2.1 Echoing concerns prevalent in the literature

The discussion ranged across many of the concerns already discussed in the literature, as presented in Sections 3, 4 and 5 above.

The lack of regulation or legislation covering the development, trialling or use of algorithmic analyses as discussed in Section 3 was a prominent concern. The discussants echoed the Council of Europe's (2015) recommendation to cement harmonisation between the various national jurisdictions which share data so that law, practice and culture can be brought closer together around the common standards of human rights, as mentioned in Section 4.2.4 above. Given that such legislation will have an impact on the work of the technology companies which develop the tools, they are in a prominent position to advise and comment on the reach, enforceability and impact of any new legislation.

The participants noted the unclear guidelines and legal code governing the sharing of data between public and private bodies, as was discussed above in Section 3.2. The growing quantity of data available will only make the need for clarification of these relationships even more pressing.

The tension that arises in balancing the investigative purpose to protect individual, national or security interests with the requirement to observe privacy rights was also highlighted by the group. These two positions were discussed earlier in Section 5.4 and the group echoed the recommendation that bodies entrusted with investigative powers be carefully regulated and held to account.

Further on the issue of accountability, the participants similarly highlighted the call for transparency in the use of the tools and the traceability of results, as discussed by The Law Society of England and Wales (2019) in Section 4.2 earlier. The gaps in current structures were considered to be unacceptable, and a minimal technical measure is the inclusion of user logging files to facilitate the tracing of past access.

A huge area of concern in the literature is bias, both arising from training data and from 'by design' from the choices made at the development stage, as summarised in Section 4.1 above. The group too recognised these dangers and recommended regular reviews to compare dataset outputs over time and the institution of flexibility to add categories or criteria of input variables to aid analysis.

Earlier in Section 5, the issues related to public trust, understanding and confidence were discussed. The group rehearsed them, noting the importance of building trust and awareness through clear communication and simply worded explanations. They also saw the value of an independent oversight body with links to the public to maintain transparency.

6.2.2 Emerging concerns

The group identified specific areas which have received less prominence in the literature to date. Of particular note was the emphasis on the special considerations needed for younger data subjects, namely people under the age of 18. Children's rights issues need close attention to ensure particular protections are in place to respect their individual rights. Safe-guarding, mental health, emotional and well-being issues have different properties for children and need to be considered in light of protecting their rights in the development, trialling and use of analytical tools. Thus, the development of a rights-based framework and mindset for the use of such tools needs to incorporate the diversity of all potential data subjects.

The participants were keen to point out the need to establish the accuracy levels of the tools with good explanations of how to interpret them. An independent evaluator was recommended to ensure distance from any conflict in interests in the tool's success and uptake.

The as-yet unknown, unintended consequences of the use of the tools was another topic elicited by the group. For example, the improper use of the tools or incidents of false positive results are some

foreseeable unintended uses of the tools and can be mediated in advance, but supervision is needed to manage unforeseeable consequences of their use.

The skillset of the personnel at LEAs to manage the fine-tuning of tools when new input data are tested or when revisions to the tool's code are needed was mentioned as a concern by the discussants. This relates to a wider issue of the after-care for ASGARD tools once the project is finished.

Related to privacy issues, the group highlighted specific techniques that could bolster anonymisation to preserve privacy, such as the removal of identifiable data; minimizing stored data; data aggregation, followed by a process to merge or remove resulting groups; identification and removal of non-key information. In situations where it is possible to control access to anonymised data, other measures can be taken to enforce anonymisations, such as synthesizing additional fake data or hiding results when aggregate queries results do not have enough elements. Further measures could be to limit access and avoid cross-dataset queries, and adding process and training to discourage attempts to de-anonymise.

Related to concerns around data retention, the discussants noted that it is not clear how the deletion of data is confirmed and communicated to the public in the context of the police use of data. The tension between the obligation to delete data and the demand to maintain national facial matching databases was also highlighted. Thus, any automated deletion capability may have repercussions for the work of law enforcement agencies, and there is also no clarity on how the retention of data is managed when they have been shared to third party.

The group also pointed to the anomaly over the status of the face as a non-biometric measure when it is as distinctive and individual as DNA or fingerprints.

Finally, the group highlighted the lack of a common and overarching ethical, legal, privacy and societal framework across H2020 projects, leading to divergent practice on different projects. They recommended a separately funded project that would provide guidance and oversight to multiple project to create coherence.

7. Conclusion

7.1 Summary

In this document we have provided an analysis of the societal impact of the introduction of the ASGARD toolset into the work of LEAs.

In section 2, it was concluded that policy makers, regulators and stakeholders all have an important role in updating legal frameworks, promoting Big Data practices, and developing and incorporating tools into the Big Data design and practice that address societal concerns.

In section 3, it was concluded that data protection legislation is highly complex and nuanced, and at present law enforcement agencies practitioners have limited sources of accessible and practical guidance on what constitutes the appropriate use of data.

In section 4, it was concluded that the primary concern with forensic analysis in ASGARD is with data input and human decision making rather than with the algorithmic analysis itself.

In section 5, specific legal questions on the remit and oversight of the LEAs in many countries was examined and it was concluded that open societies have to protect themselves, but the parts of the state entrusted with significant powers must be carefully regulated and held to a high level of accountability.

In section 6, the workshop that provided current commentary and reflection on the societal, ethical, privacy and legal aspects of the ASGARD toolset was discussed. The output of the discussion parallels many of the concerns discussed in the literature in the previous sections, reflecting the current trend in the concerns around the use of AI and machine-learning analytical tools in the police context. There were, in addition, some concerns and different emphases placed on concerns that are less well reported in the literature. These concerns were around use of data related to minors; evaluation and accuracy; skillsets; and anonymisation.

Finally, it is worth highlighting that participants at the workshop all had a concern about the lack of a common and overarching ethical, legal, privacy and societal framework across H2020 projects, leading to divergent practice on different projects.

7.2 Evaluation

The paper reviewed recent literature on the social impact of Big Data. The BYTE project showed that in order to address discrimination, equality and trust, privacy-by-design methods should be extended to anti-discrimination-by-design and analogous approaches, and transparency and new accountability frameworks need to be based both on legislation and on a data protection framework. Recent literature on LEAs and the challenges of Big Data evidences that data protection legislation is highly complex and nuanced, and at present law enforcement agencies practitioners have limited sources of accessible and practical guidance on what constitutes the appropriate use of data. There is a push/pull tension between the need for a secure society and civil liberties. The Snowden disclosures show how such data collection can be undertaken without public awareness or consent. Such awareness and consent are crucial, as are robust oversight mechanisms to reassure the public that capabilities are not being misused or abused. There are further concerns about what happens to an individual's data after it has been collected, in particular the circumstances in which this data is interrogated and analysed and for how long data is kept.

The results of algorithms in the investigative function of LEAs is liable to bias if the training data or input data maintain skewed data unless attempts to control for it, monitor or re-balance the data are

made. Training data is not the only source of bias as it may also creep in from value-laden decisions during the development stage. Recommendations to root out or minimise bias will be forthcoming from the consultation undertaken by RUSI for the UK's Centre for Data Ethics and Innovation. Undoubtedly, they will include recommendations on accountability structures and regulation of algorithmic analytics; transparency; stronger legislation and regular equality impact assessments. The challenge of harmonising law, practice and culture between different jurisdictions which share data also needs addressing.

The importance of public confidence and trust is demonstrated in the literature. There is variation in public awareness of how data are collected and used, and in opinion over the use of personal data. Yet, there is a need for public support for legitimate state law enforcement and intelligence work in principle. Again, the oversight frameworks for governing the use of data in the LEA context needs to be practical and credible to garner public support.

In order to seek feedback on the societal impact of the ASGARD toolset, a workshop was held to take participants through a staged process, which included a demonstration of the typical use of the toolset and the legal/ethical framework, and facilitated discussion. The commentary provided by the participants from a wide range of professional backgrounds underscored many of the existing concerns reported in the literature and also highlighted emerging concerns that are less well reported.

7.3 Future Work

The findings of this analysis of the societal impact of the introduction of the ASGARD toolset into the operational work of LEAs suggests that work still needs to be done within and beyond the project to take account of the concerns that exist. Many are beyond the remit of ASGARD, but some sit well within reach of the project partners. These include: the accuracy of the tools needs to be assessed, user logging files for traceability are needed, recommendations to LEAs on possible bias in outputs should be developed, sufficient expertise needs to be available after the project ends to support revisions or fine-tuning of tools, simply worded explanations of what the tools do are needed to promote public confidence and support, and these should include clear indications of data deletion protocols, and an audit of where possible threats to children may arise in the tools' processing needs to be undertaken.

8. References

Alpaydin, E. (2016) *Machine Learning*. MIT Press, Mass.

Babuta, A. (2017) *Big Data and Policing - An Assessment of Law Enforcement Requirements, Expectations and Priorities RUSI Occasional Paper*. https://rusi.org/sites/default/files/rusi-bigdata-press-2017.pdf

Babuta, A. and Oswald, M. (2019) *Data analytics and algorithmic bias in policing,* Royal United Services Institute for Defence and Security Studies. Available at: https://rusi.org/publication/briefing-papers/data-analytics-and-algorithmic-bias-policing.

Council of Europe, Commissioner for Human Rights (2015) *Democratic and Effective Oversight of National Security Services*, Issue Paper, Council of Europe.

Cuquet, M., Vega-Gorgojob, G., Lammerantc, H., Finnd, R. and Hassane, U. (2016) *Societal impacts of Big Data: challenges and opportunities in Europe* https://arxiv.org/ftp/arxiv/papers/1704/1704.03361.pdf

House of Commons Science and Technology Committee [HCSTC] (2018) *Algorithms in decision-making. Fourth Report of* Session 2017-19. Available at: https://www.parliament.uk/business/committees/committees-a-z/commons-select/science-and-technology-committee/inquiries/parliament-2017/algorithms-in-decision-making-17-19/. Para 37-38.

Ipsos MORI (2014) *Understanding Society: The Power and Perils of Data*. Available at: https://www.ipsos.com/ipsos-mori/en-uk/understanding-society-power-and-perils-data

The Law Society of England and Wales (2019) *Algorithms in the criminal justice system – A report by the Law Society Commission on the Use of Algorithms in the Justice System*. Available at: https://www.lawsociety.org.uk/support-services/research-trends/algorithm-use-in-the-criminal-justice-system-report/

Michael, K. and Millar, K.W. (2017) 'Big Data: New Opportunities and New Challenges,' *Computer,* Vol 46, issue 6.

Mittelstadt, B.D., Allo, P., Taddeo, M., Wachter, S. and Floridi, L. (2016) 'The ethics of algorithms: Mapping the debate,' *Big Data & Society*, July – December, pp1 – 21.

O'Neil, C. (2016) *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Penguin Books. p86.

Oswald, M. and Grace, J. (2016) 'Intelligence, policing and the use of algorithmic analysis: a freedom of information-based study,' *Journal of Information Rights, Policy and Practice*, 1(1).

Oswald, M., Grace, J., Urwin, S. and Barnes, G.C. (2018) 'Algorithmic risk assessment policing models: lessons from the Durham HART model and 'experimental' proportionality,' *Information & Communications Technology Law,* 27(2), pp223-250. p235

Richards, N.M. and King, J.H. (2014) 'Big Data Ethics,' *Wake Forest Law Review,* 49, pp393-432.

Royal United Services Institute for Defence and Security Studies (2015) *A Democratic Licence to Operate Report of the Independent Surveillance Review*, Whitehall Report 2-15, July 2015. https://rusi.org/sites/default/files/20150714_whr_2-15_a_democratic_licence_to_operate.pdf

The Royal Academy of Engineering (2018) *Written evidence submitted by The Royal Academy of Engineering,* Available at: http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/science-and-technology-committee/algorithms-in-decisionmaking/written/69070.pdf)

Ruggieri, S. & Pedreschi, D. & Turini, F. (2010) *DCUBE: Discrimination discovery in databases*. Proceedings of the ACM SIGMOD International Conference on Management of Data. 1127-1130. 10.1145/1807167.1807298.

YouGov/*Sunday Times*, 'Survey Results', 15–16 January 2015, https://d25d2506sfb94s.cloudfront.net/cumulus_uploads/document/wt26kxdn72/YG-Archive-Pol-Sunday-Timesresults-160115.pdf

ANNEX I. **ASGARD GRANT AGREEMENT: SOCIETAL IMPACT (PP 147-150)**

5.2. SOCIETAL IMPACT

The impact of security developments on society can be complex. This is because some measures which are designed to enhance security can be at variance with Europe-wide societal values such as respect for privacy, equality, and the rule of law, which must always be upheld. Issues of ethics, legality, and privacy are therefore at the core of ASGARD, particularly in terms of maintaining a measure of proportionality between national and citizen security and civic rights and how this is managed in practical terms. This explains why in the administrative proposal forms (Part A), we replied 'yes' to 9 of the 10 sub-questions in the Societal Impact Table, including 'yes' to ASGARD research both benefiting but also potentially having a negative impact on society; our reasoning is detailed further below.

5.2.1. DOES YOUR RESEARCH MEET THE NEED OF SOCIETY?

Does the proposed research address documented societal security need(s) (e.g. life, liberty, health, employment, property, environment, values)?

The EU is currently faced with a diversity of complex human-made cross-border threats, many with a technological component including, for example, regional instability, violent extremism and terrorism, organised crime, cyber offenses, and the combination of these. Such threats pose very real and direct threats to the lives of individuals (in terms of life and liberty), to their employment prospects and property rights, to the broader environment, and to citizens' overall sense of safety. At the same time, much of society's 'sense' of security is based on accepting and adhering to the core values of human dignity, freedom, democracy, equality, the rule of law and respect for human and minority rights, values enshrined in the European Convention on Human Rights (ECHR). ASGARD has been designed to offer additional technology-based protection to citizens against acts of terrorism and organised crime while remaining cognisant of these other core values and seeks to find a 'socially acceptable balance' between them. It recognises, further, that while technology can be part of an effective response to these threats, it can only do so in conjunction with an understanding of human factors and organisational processes (i.e. societal issues and structures), including the different experiences and cultures of EU member states and agencies. Does the research output meet these needs? Will this be demonstrated? Will the level of societal acceptance be assessed? Demonstrating how the toolset helps meet societal needs is both simple and complicated. It is simple in the sense that a reduction in the number of, for example, terrorist incidents will be an expected outcome. However, the more tangible material cost is more difficult to quantify since it is complicated to estimate the cost of a

terrorist attack that has been averted and the cost of lives saved especially as against increased fears in the wake of the Snowden revelations around seemingly unchecked mass surveillance, including by EU member states, in the name of counter-terrorism. It is unlikely, in other words, that a toolset such as ASGARD will be unconditionally welcomed by citizens despite their acknowledgement of the threats it seeks to counter. ASGARD will thus demonstrate improved operational and situational awareness for LEAs whilst at the same time paying close attention to citizens' legitimate concerns and associated levels of societal acceptance.

The latter will be assessed in two main ways. Firstly, the ASGARD proposal includes an Ethical and Societal Impact Review Board independent of the consortium, which will meet regularly to ensure that the methods and results of the project meet expected standards of methodology, rigour, and conduct in a manner acceptable by citizens. Secondly, the participants involved in Work Package 12 will develop a societal acceptance attitudinal survey to test the findings and outcomes of the proposal. This will assure that civic rights, especially with respect to citizens' right to privacy, are at the core of this WP.

Does the research address threats to society (e.g. crime, terrorism, pandemic, natural and man-made disasters etc.)? Does the proposed research address in an appropriate way these threats?

The ASGARD research proposal addresses explicitly the major current threats, risks and vulnerabilities documented in The European Agenda on Security (2015). The primary research output of ASGARD will be the creation of a toolset, complimentary to tools already in use and available in the market, which makes a significant leap forward in the technological competencies of Law Enforcement Agencies. Instead of focusing primarily on the analysis of information for the purposes of generating evidence with legal veracity, ASGARD will work on the generation of Primary Intelligence in order to ensure that investigators can manage and exploit the increasing volumes of mixed media data connected with on-going investigations and with crime trend identification and prevention.

5.2.2. DOES YOUR RESEARCH BENEFIT SOCIETY?

Does society as a whole benefit from the proposed research?

All segments of society will benefit from this project, either directly or indirectly. For instance, the general public will benefit from a likely reduction in the number of terrorist and cyber incidents whilst LEAs will be aided in overcoming a significant management challenge and thus more efficiently obtaining situational and operational awareness. Issues of 'ethics and justice' are at the core of the ASGARD proposal particularly in terms of maintaining a measure of proportionality between the right to security and civic rights and how this is managed in practical terms. The ASGARD proposal considers the factors that cause citizens' feelings of security and insecurity as part of its efforts in ascertaining levels of societal acceptance within the project (WP12). The 'culture of public users' will be considered

in terms of the tools to be used for trawling, bias in analysis, any possibility of misuse by the LEAs, any cultural differences between LEAs, and the general public perception of the countries security forces.

5.2.3. DOES YOUR RESEARCH HAVE NEGATIVE IMPACT ON SOCIETY?

Are there other European societal values that are enhanced by the proposed research (e.g. public accountability and transparency; strengthened community involvement; human dignity; good governance; social and territorial cohesion; sustainable development etc.)?

The expected outcome of LEAs using the toolset should be the ability to work in shortened periods of time to intervene and prevent acts of terrorism and organised crime. The proposed research will, further, contribute to the enhancement of public accountability within the LEAs. It could also contribute to the enhancement of community engagement, social and territorial cohesion and the minimisation of inequalities in terms of the abilities of LEAs to engage in data mining activities as part of their activities in the prevention of terrorism and terrorist incidents. In addition, good governance can be enhanced by giving better, faster and more transparent access to raw information on crime. This output will, of course, be in line with the need to protect people's security as well as the need to enshrine and protect people's civil rights and civil liberties and so the issue of privacy and security mapped against the toolset.

If implemented, could the research have a negative impact on the rights and values enshrined in the Treaties (e.g. freedom of association, freedom of expression, protection of personal dignity, privacy and data protection)?

The overall objective of WP12 is to ensure that the toolset to be developed will take into account any social, privacy, ethical, legal, economic and regulatory implications both at its development stage and at its testing stage; to anticipate the illegitimate use of the toolset in its operational phase; to highlight the regulatory mechanisms then needed to ensure democratic accountability and prevent misuse. Successful outcomes will result in the detection of great number of individuals and groups of people associated with terrorism. In the longer term, the number of detected people may be expected to decrease as the technology success will act as a dissuader. Detection of greater numbers of individuals involved in terrorism may in the short term give rise to increased public concern about their individual safety and their perceptions of safety. On the one hand, ASGARD addresses and deals with the prediction and prevention of terrorist incidents which have the potential to result in extended negative outcomes for broader society. On the other hand, ASGARD will contribute to embed a value-based toolset which in turn would lead to societies that are better placed to promote and support democracy, human rights and the rule of law which will lead to citizens having a greater sense of their own security in a changing world. Large-scale data gathering exercises have a degree of risk – securing authorized access only to sensitive data; abuse of profiling; lack of trained personnel with experience to interpret data in line with the correct guidelines on personal dignity and civil liberties.

Therefore, any potentially negative impacts of the ASGARD research proposal would be as a result of the absence or lack of technical safeguards, of end-user knowledge and of national legal safeguards ensuring that any limitations on civil liberties and rights generated by the envisioned toolset, particularly in relation to issues of privacy. Thus, the protection of a person's right to privacy and broader data protection will be carefully addressed throughout the project to ensure that all legal and ethical constraints are met.

If implemented, could the research impact disproportionately upon specific groups or unduly discriminate against them?

The ASGARD proposal will focus on the prevention of terrorism and terrorist incidents. Successful outcomes will result in the detection of individuals and groups of people associated with terrorism. In the longer term, the number of detected people may be expected to decrease as the technology success will act as a dissuader. Detection of greater numbers of individuals involved in terrorism may in the short term give rise to increased public concern about their individual safety and their perceptions of safety. In terms of a disproportionate impact against a specific group or cohort, there could theoretically be some misuse of, or arbitrary decision made following automatic criteria recognition, especially face recognition, sentiments analysis or opinion mining; an arbitrariness of the results of database interlinkage; and an absence / lack of legal safeguards surrounding the use of the envisioned system by LEA. Moreover, semantic analysis, image, text and audio analysis might discriminate against certain groups if trained on prejudiced data sets.

Will specific measures be taken to ensure that the research outcomes comply with the European Charter of Fundamental Rights and to mitigate against any of the negative impacts described above?

To ensure compliance with the European Charter of Fundamental Rights and mitigate the negative impacts described previously, the proposed system requirements and specifications of ASGARD will take into account individual rights, principles of proportionality, human dignity and non-discrimination not only in relation to their legal, formal expression but also regarding the ethical aspects related to them. Within the time frame of the proposed project a new European Directive on Data Protection will enter into force and begin its process of transposition into national law.

The ASGARD proposal acknowledges the need to strive to find the best balance between the valid ethical concerns of society on the use of Data Mining and Information Analysis through dedicated technology and the practical aspects of Law Enforcement which ultimately strive to protect and serve society. The project involves a broad range of parties with experience in legal, ethical and privacy issues as well as LEAs and research groups. Bringing together this expertise will mitigate many of the privacy risks previously discussed. Training has also been made a key consideration. Ultimately, ASGARD is based upon the premise that adherence to privacy and ethical norms helps enhancing European values, which lead to societies that better promote and support democracy, human rights and the rule of law. Thus, citizens have a greater sense of their own security in a changing world.

ANNEX II. SUMMARY OF SOCIETAL IMPACT WORKSHOP

Summary

Expert discussion on law enforcement use of machine-learning analytical tools and their impact on privacy rights and data protection

On 18 September 2019, at Dublin City University, 24 participants took part in a discussion on the topic of the use of machine-learning and other analytical tools by law enforcement agencies (LEAs) and its impact on privacy and human rights. The participants came from a range of professional backgrounds: human, child and privacy rights organisations, technology companies working on ASGARD, LEA partners of ASGARD and from other forces, legal professionals and academics. The list of participants is given at the end of this document.

Through brainstorming and focused discussion, the participants addressed the following questions:

> What are the privacy, data protection and ethical concerns about the development, introduction and use of AI and machine-learning analytical tools intended for use in criminal investigations, criminal behaviour forecasting and surveillance.
>
> What are the concerns about the use of the tools by law enforcement agencies? What can go wrong?
>
> What needs to be in place to safeguard against things going wrong? By whom?
>
> What should happen if things do go wrong? Who should do it?

The following issues of concern and recommendations were raised for consideration. The issues have been grouped under six thematic headings.

Regulation and guidelines on use of tools

a. Lack of legislation regulating algorithms

There is a lack of legislation in several EU countries regulating the development and use of algorithms in general. This absence needs to be addressed. However, there is a fear that regulation may stifle innovation.

b. Lack of a common ethical framework across jurisdictions on how LEAs use machine-learning and other analytical tools in their work

There are different approaches in different member states and there is a case for a common ethical framework which should adhere to the human rights framework guided by privacy rights and data protection regulation. Jurisdictional experts should review the tools before deployment.

User agreements and other protections are needed regarding the use of the tools, for example licences, which set out guidelines on their use, including ethical and human rights considerations.

User protocols should flag any restrictions on the tool's use and reminders of the legal framework of originator agency. They should also include where and how information can be used, by whom. The tools should include prompts to guide proper use.

c. Develop a human rights mindset across development and deployment of such tools

There should be awareness, understanding and appreciation among all groups of experts who engage with the development and regulation of tools of the different perspectives held by developers, users and legal overseers. For example, technologists need to understand legal issues; legal experts need to understand LEA needs for technology and data.

Underpinning this is that developers, legal commentators and LEAs groups should have a basic understanding and appreciation of fundamental rights and freedoms which any tool developed should respect. Use the European Convention on Human Rights as a basis.

This would help to create a culture of respect for the need to balance investigative demands, technological capabilities and privacy when developing tools which will have an impact on people's personal lives, data and rights.

d. Develop a mindset to protect the rights of children

Related to the previous point are the specific protections related to children. The use of analytical tools should be respectful of children's rights as a subset of wider human rights.

There needs to be particular measures in place when data relating to a person under the age of 18 are flagged, for example, consider a different process in order to respect the rights of the child. These measures need to be baked in at the time of the development of the tools. Of importance are safeguarding considerations, mental health, emotional and well-being issues.

Consideration should be given to the use of training data that includes children. Consideration should be given to whether greater protections to avoid particular risks need to be in place. As does consideration for the inclusion of specific rules in relation to children's data.

e. Different regulatory control for data for intelligence or evidence

In the LEA context, there is a difference between data that are used for investigative intelligence and digital forensics. Digital forensics is subject to a chain of evidence and can be regulated by minimum standards. The prosecution conducts the forensic analysis and the defendant has to rely on it without necessarily knowing how it has been obtained. For example in the UK, legal regulation is provided by Police and Criminal Evidence Act 1984 and Computer Misuse Act 1990.

Evidence that is required to be disclosed to the defendant may require technical know-how to do this.

f. Unclear legal restrictions on public-private data sharing.

Clear protocols or rules, guided by legal, ethical and moral considerations, are needed on when, how, who, why and what data can be shared between private, public, national bodies, both internally and with other countries.

g. Tension between privacy and other human rights and national security

Maintaining the balance between privacy and data protection and the protection of national interests and national security is an on-going process. Privacy rights cannot be pushed to one side in the single-minded pursuit of national interests but have to be given due respect.

Privacy issues

a. Anonymisation is not a fail-safe

Anonymisation is not necessarily an assurance of preserving privacy. Some best practices can be adopted to minimize its weakness, such as removing identifiable data; minimizing stored data, removing non-key information; data aggregation, followed by a process to merge or remove resulting groups with not enough cardinality. Whenever it is possible to control access to anonymised data, some measures can also be taken to enforce anonymization, such as synthesizing additional fake data or simply hiding results, when queries results do not have enough elements. As additional measures, access can be limited and cross-dataset queries avoided as well as add process and training to discourage attempts to de-anonymise.

b. Data retention

There is an obligation to delete data, but how is this confirmed and communicated to the public in the context of the police use of data? How does this obligation tally with the national facial matching database? Automated deletion capability can be instituted but this may have repercussions for law enforcement. The future quantity of available data needs to be considered. Also, how is the retention of data managed when they have been shared to third party.

c. Encryption of data

The use of encryption may be a problem in the long-term for LEAs. Yet, the public like and trust encryption.

d. What constitutes biometric data?

There is an anomaly over the status of the face as a biometric measure in need of protections in the accorded to DNA and fingerprints.

Consequences and accuracy

a. Oversight and accountability mechanisms

To track when, by whom and for what purpose the tools are used, user log files should be incorporated into the tools or the user interface.

The capability to challenge a result should be available in case there is a legal challenge to a result that was derived fully or partially by automated means.

To facilitate understanding of the workings of the tools, there needs to be clear explanations of how they work written for lay-people. There needs to be clear guidance and accessible mechanisms for 'policing the police.'

b. Bias

'Bias-by-design' can arise from training data and from the design of the tool. It is necessary to institute external and independent testing and checking to ensure bias does not beget bias. This could be through comparing dataset outputs, including over time as input data increase. Input datasets and training data should reflect diverse populations. Also, there should be flexibility to add categories or criteria of input variables, and to correct or reset AI to prevent bias. Regular reviews, trend analyses or other reports on the use and outputs of the tools should be instituted.

The capability of LEAs to undertake revisions of a tool's code or functionality needs to be assessed.

c. Accuracy of results/output

What constitutes high or low accuracy needs to be determined, with adequate explanation of how accuracy scores are derived and how to interpret them. If accuracy needs to be improved then the human users may need to know what they mean and an independent evaluation of the tools may be required to establish their accuracy.

d. Unintended consequences and uses

If a tool is used for state purposes that invades the privacy of individuals in a *bona fide* way and is then abused (for example, to prosecute lesser morality offences) or is over-used, there is a danger that the court will rule it as an abuse and the tool could be lost. Tools should be restricted to specific purposes, for example for 'seven years and up' offences, or restricted so that the information analysed is consistent with the investigation's purpose or other criminality uncovered. Fishing capabilities should be prevented.

Mitigation methods for incidents of false positive results or the involvement of individuals who are not associated with criminal aspects of an investigation need to be developed for LEAs.

Public trust and awareness

a. Build public trust

There is a presumption in the public eye that LEAs are doing something unlawful. This needs to be addressed to build trust. A public forum for presenting, monitoring and challenging the tools is needed. It will help to improve trust and confidence and build the legitimacy of the tools. It will be able to build awareness of the tools, possibly at the outset of the project. It should have independent authority with annual evaluations and direct links to the public.

Transparency is needed with regards to why certain information is not shared, the failure rates and efficiency statistics of such tools and the legal basis or using them.

While there is a tendency among some populations not to be concerned about their data, there is a good level of understanding of GDPR.

b. Communicating with the public about the tools and their deployment

Any informed discussion needs to use non-specialist language, relatable examples, directed to the purpose or need of the tools, focused on safeguards and remedies. Discussion should be guided by risks, costs and eventual use.

A representative body could be established, such as the citizens' assembly model, to engage the public about the tools. It could have age limits or it could be independent.

c. Defining responsibilities

For the public to understand what the tools do and the responsibilities they place on LEAs, there needs to be clear statements about their intended use and how results from them are used in decision-making. This includes any issues that may arise over the autonomy and responsibility of the human operator in interpreting and adopting results.

d. Independent body to oversee

There needs to be intermediary reviews and evaluations. The decision-making power needs to be decided but it may be dependent on privacy intrusions or sensitivity of the data.

Assigning responsibilities

a. A machine is never responsible

The mantra 'a machine is never responsible' should be embedded in any LEA operational setting. Intelligence foresight tools require the human factor to conduct checks or balances, for example, through supervision or oversight by a human. There is no need for regular checks as it is immediately clear what is working and what is not working. Human intervention is particularly important for pipeline tools and a filter ensuring human intervention could be built in.

b. Where does the responsibility lie for the accuracy of results?

Between the developer of a tool and the user, is it clear where the responsibility for ensuring the highest level of accuracy lies?

c. Fine-tuning tools

Personal non-anonymised data are used in training datasets to train tools. The tool is then tested on 'real' data and may need to be fine-tuned. There is an issue over the capacity of the end-user to manage this process of fine-tuning or retraining.

Project-specific issues

a. Operational issues after project-end

Related to the point above, it needs to be established to whom any problems with toolkits developed by ASGARD are addressed. Also, have measures been considered about how to future-proof the tools and to preserve the legacy of ASGARD? Will the network continue after the funding period has finished?

b. Wider H2020 issues

Intellectual Property and other learning from H2020 projects, both former and current projects, have resulted in a duplication of effort with consequent diminishing returns. For example, there needs to be an overarching ethical, legal, privacy, societal framework across common H2020 projects. This could be formed as a project within a project with separate funding and oversight of multiple projects for coherence.

Purpose of this summary

The summary will be incorporated into ASGARD deliverable on societal impact for the Societal, Ethical, Legal and Privacy work package. Additionally, participants in the discussion are free to circulate or refer to it as they wish.

Participants

1	Mr Alexander Babuta	Royal United Services Institute for Defence and Security Studies, UK
2	Col. Armin Vogl	Federal Ministry for Internal Affairs, Austria
3	Mr Bernardo Pacheco	INOV, Portugal
4	Ms Christine Andreeva	Dublin City University
5	Dr Fredrik Johansson	Swedish Defence Research Agency
6	Dr Gary Ellis	The University of Guelph-Humber, Toronto
7	Mr Gordon Doyle	IBM
8	Ms Jane Hollway	Ulster University
9	Mr John Barry	INTERPOL
10	Dr Lucy Royal-Dawson	Ulster University
11	Ms Margaret Gallagher	National Society for the Prevention of Cruelty to Children, NI
12	Mr Mark Potkewitz	Ulster University
13	Prof Martin Haran	Ulster University
14	Prof Maura Conway	Dublin City University
15	Advocate Nery Ramati	Dublin City University
16	Dr Omar Nibouche	Ulster University
17	Dr Owen Corrigan	Dublin City University
18	Ms Renate Samson	The Open Data Institute
19	Prof Ruth Fee	Ulster University
20	Mr Stephan Veigl	Austrian Institute of Technology GmbH
21	Mr Stephen Murray	Ulster University
22	Dr Suzanne Little	Dublin City University
23	Mr Tony Fisher	Human Rights Committee of the Law Society of England & Wales / Fisher Jones & Greenwood Solicitors
24	Anon - name and organisation withheld upon request	