

Regulatory Sandbox Final Report: Tonic Analytics

A summary of Tonic Analytics participation in the ICO's
Regulatory Sandbox Beta

Date: February 2021

ico.

Information Commissioner's Office

Contents

1. Introduction	3
2. Executive summary	4
3. Product description.....	5
4. Key data protection considerations	9
5. Ending statement	20

1. Introduction

- 1.1 The ICO introduced the Sandbox service to support organisations who are developing products and/or services that use personal data in innovative and safe ways and where such products and/or services deliver a potential public benefit. In order to develop the Sandbox, the ICO launched the Sandbox as a beta phase, for an initial group of participant organisations during 2019 – 2020. The beta phase provided a free, professional, fully functioning service for ten organisations, of varying types and sizes, across a number of sectors.
- 1.2 Organisations who were selected for participation in the Sandbox beta phase have had the opportunity to engage with us; draw upon our expertise and receive our advice on mitigating risks and implementing 'data protection by design' into their product or service, whilst ensuring that appropriate protections and safeguards are in place. Tonic Analytics' (in relation to their Galileo Programme) was one of the candidates selected for participation in the Sandbox beta phase.
- 1.3 Tonic Analytics acted as a central coordinating point (on behalf of the Police and Highways England) for work involving researching the possible viability of using up to date data analysis to improve road safety, reduce the economic impact of road incidents and help address criminal behaviours on the roads. This work was entitled the 'Galileo Programme'.
- 1.4 The original planned work in the ICO Sandbox had included operational pilots which would have used live personal data in a limited fashion. However, due to the impact of COVID-19 this work could not be undertaken within the available timeframe due to the significant impact on stakeholders, therefore resulting in delays to the work of the programme.
- 1.5 Therefore, the work conducted over the course of the ICO Sandbox participation was limited to specific non-operational research pilots either using pseudonymised data or using no personal data at all, and all outputs were presented in an aggregated, anonymised format and with no action being taken against individuals as result of the pilots. This means that no identifying data was passed back to police forces on the basis of the research.

2. Executive summary

- 2.1 The Galileo Programme, under the direction of the Police and Highways England, aims to enable intelligence led actions that aim to reduce the number of fatally and seriously injured road users in the UK and reduce the amount of crime associated with use of the UK's roads.
- 2.2 The application of advanced analytics on big data sets is key to the programme aims of reducing the number of serious collisions and crimes on the roads. This intelligence is intended to support more efficient data sharing between public and private sector organisations, enable resource to be targeted effectively, and deliver coordinated actions in the public interest.
- 2.3 The work conducted during the Sandbox participation was limited to two research pilots, known as the 'Speed' and 'Intelligence, Tasking, Admin and Performance Unit "(ITAP)" pilots. Both pilots used pseudonymised personal data. A further pilot conducted at the same time contained no personal data at all and therefore is not in scope of data protection law. Although the processing associated with the third pilot did not fall within the scope of the data protection requirements, it enriched the overall context of the ICO's understanding of the Galileo Programme and was therefore still included in the Sandbox plan.
- 2.4 The Galileo Programme has made significant progress in understanding what best practice looks like in terms of the use of data and its own data protection requirements and has achieved this by using pseudonymised data in a way which reduces risk to the data subject. This work will form an instructive basis on which an operational model can be built, which could be of a material benefit to the police and other public authorities and improve the safety of the public at large.
- 2.5 The benefits of serious harm reduction must be balanced against both the rights and freedoms of the data subjects and their reasonable expectations to maintain a level of privacy as they travel across the road network. Full and proper consideration should be given as to any other less invasive ways the same aims could be achieved. Before the Police Forces process any kind of personal data in relation to the Galileo Program, they must ensure that it is based upon a specific, legitimate aim,

with a basis in law under part 3 of the DPA 2018. The considerations of the involved in balancing between the proposed benefit and invasiveness of the processing may mean the approach is be less likely to be permissible if the programme is used for lower level policing involving infractions such as expired road taxation or road toll evasion. However, it is also recognised that the Programme has both identified new links, and validated existing evidence, that strong relationships exist between low level offending behaviour and more serious criminal behaviour. One of the core objectives of the Galileo Programme is to explore ways that roads policing activity can assist in the detection, prevention and prosecution of all forms of criminal behaviour, including serious and organised crime.

3. Product description

- 3.1 Tonic Analytics are responsible for developing a platform and programme of functionality entitled "Galileo", on behalf of and funded by Highways England and the Police, to help tackle several areas of vehicle crime and legal non-compliance on the roads.
- 3.2 Central to the proper functioning and benefit realisation of the programme is the application of a big data processing platform, which will use analytics tools to provide non-personally identifiable insights into factors associated with driver behaviour. This technology has been repurposed from Tonic Analytics' existing aerospace applications. The insights generated will be based on the processing of pseudonymised data but outputs are aggregated to a level where they are no longer identifiable. Data will be sourced from multiple public sector bodies under a range of different data sharing agreements. The processing aims to determine issues around current assumptions relating to on road vehicles such as links between types of crime and also has the ability to conduct exploratory exercises in order to indicate areas which have not been previously considered by the Police or Highways England. Functionality to support the Police in focusing on more serious and risky behaviours on the roads, includes the creation vehicle profiles.
- 3.3 The core approach under this project, with the objective of delivering better data driven intelligence, involves the use of cloud based computing, data analytics technology (including artificial intelligence) and the building of a comprehensive

blended big data set that models the road eco-system, which is referred to as a 'Digital Twin'. The Digital Twin data model incorporates extensive data relating to:

- vehicles using the road network;
- the behaviours and actions of vehicle keepers and drivers;
- the road network and traffic conditions; and
- environmental and temporal factors.

- 3.4 After the conclusion of the Sandbox participation, Galileo plans to offer operational services to police forces based upon the same or similar data sets. Outputs from data processing will be used to support operational decisions enabling police forces to deliver safe law enforcement outcomes. Analytics will provide both predictive and responsive intelligence, with overall situational awareness being provided through business intelligence analytics.
- 3.5 The two pilots concerned with the processing of personal data during the course of ICO Sandbox participation are outlined below:
- 3.6 Speed Pilot The key objective was to provide data driven insights that can better inform future intervention methods aimed at addressing speed related offences, such that resources are more efficiently and effectively utilised in delivering good outcomes for the public.
- 3.7 The pilot sought to better inform compliance and enforcement activities conducted by the roads policing community. It was focused at specific locations, predominantly on the Strategic Road Network (SRN) (i.e. the motorways and major A roads), which is better instrumented than other roads to provide relevant data for analysis due to the nature of available data provision and camera technology. Data relating to offending speed behaviour included both speeds in excess of a prescribed limit and also speeds considered too great for prevailing conditions, typically reported as dangerous, careless or antisocial driving/riding.

3.8 The pilot examined traffic flow and speed camera data over a period of several months, blended with existing data within the Digital Twin environment. The research specifically sought to identify the relationships between compliance and offending in the context of the speed of the vehicle, when compared with other driving and road user behaviours, as well as journey and vehicle types, as can be anonymously identified within the data sets comprising the Digital Twin.

ITAP pilot

3.9 This pilot combined a broad range of data sets relevant to law enforcement and road safety. For the purposes of:

- Researching the potential and value of enhanced intelligence products (eg by blending other data sets with their Automatic Number Plate Recognition hotlist data, which often relates to untested and single-source intelligence such as an anonymous Crime Stoppers report); and
- Performing research aimed at generating new insights into links between different types of offending behaviour, as well as to different environmental factors such as driving conditions (daylight, darkness, precipitation, etc).

3.10 Additionally, the pilot would examine the current ITAP tasking and coordination function (including performance management) and, working with the roads policing unit, seek to demonstrate how this could be enhanced through further evolution of the earlier activity which demonstrated a prototype cloud hosted software application to provide end-end structured tasking and work-flow management for use by police forces. This tool specifically focuses on ensuring that all actions and interventions could be directly related to identified law enforcement problems and that the impact of those actions and interventions could be quantified.

Note

3.11 It is important to recognise that, by applying the data minimisation principle¹, the only items of directly identifiable personal data included in any Galileo data set are the Vehicle Registration Mark (VRM) and the Vehicle Identification Number (VIN). No data set, either requested or received, contains any individual's name or address. Tonic Analytics have implemented

¹ GDPR Article 5(1)c - Personal data shall be - adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed ('data minimisation');
DPA 2018 part 3 section 37 - The third data protection principle is that personal data processed for any of the law enforcement purposes must be adequate, relevant and not excessive in relation to the purpose for which it is processed.

measures, through the data protection impact assessment ('DPIA') process, to ensure that the Programme team does not come in to contact with any other data that may result in an individual being identified.

4. Key data protection considerations

The scope of the data protection considerations of the programme and identify appropriate legal framework for the processing of personal data

- 4.1 At the outset of its Sandbox participation, Tonic Analytics requested support in relation to: (i) defining the appropriate legal relationships between the parties involved in the programme in relation to the data, and (ii) a framework for the sharing of data and the cooperation of partner organisations in relation to the programme. As many of the partner organisations were public bodies with different remits, and because collecting data under different legal frameworks and the commissioning of the programme was conducted by two separate public bodies, this added to the complexity of determining the relationships. Over the course of August 2019 - March 2020 the ICO reviewed documents provided by Tonic Analytics regarding the roles

and responsibilities of the participants in the programme and the role of Tonic Analytics as the coordinating organisation. The ICO took into consideration the funding for the programme from Highways England and the Police and the ultimate aims and objectives of the processing.

- 4.2 After initial discussions with the stakeholders in the Programme, the view was taken that it was ultimately for the stakeholders themselves to decide the legal relationship between themselves. However, to comply with the Law, any data collected for law enforcement purposes under part 3 of the Data Protection Act (DPA) 2018 should only be further processed for reasons other than law enforcement if it is legally permissible to do so. It was determined that Tonic Analytics could only process data under part 3 of the DPA 2018 as a contracted processor on behalf of a competent authority (acting as controller), with Tonic Analytics meeting all other requirements specified by the controller of the personal data, with specific attention paid to those areas established as requirements as set out in law.²
- 4.3 The stakeholders in the Programme collectively concluded that the best way forward was to nominate Gloucestershire Constabulary as the responsible data controller, for maintaining contracts with Tonic Analytics. This was decided due to a prior arrangement already in place, for Gloucestershire Constabulary to act as the responsible controller for the limited

² Please see DPA 2018 Part 3 Section 59

exercises in data processing carried out during the pilot phase of the programme, ensuring that appropriate contractual agreements had been put in place between Gloucestershire Constabulary and all other partner organisations providing data for the purposes of the pilot. Tonic Analytics would then process data under instruction of Gloucestershire Constabulary and with appropriate sign off from their responsible Data Protection Officer. The ICO requested that the programme stakeholders ensure that appropriate transparency information was provided by stakeholders where they were sharing personal data with the police to ensure that the rights of the data subject could be observed across the entirety of the data lifecycle. Reviews were undertaken of all available publicly available privacy information to reflect this.

Retention of personal data for operation purposes and research purposes

- 4.4 The ICO Sandbox was asked to help support a review of Tonic Analytics' proposed retention periods, and to enrich the organisation's understanding of the applicability of any exemption from the requirement to have a clearly articulated publicly available retention period or review policy for personal data. During the course of working with the Galileo Programme it became clear in any event, that the data required for demonstrating the proof of concept, would only be needed to be held for the duration of the pilot. Tonic Analytics documented this retention period for the pilot phase and described it in the DPIA as 12 months after receipt of the data or 30 days after the completion of processing for the purposes of the research pilots as described above (as defined by the completion date of Gloucestershire Constabulary's Contract) whichever is earliest.
- 4.5 As the work progresses into the future, data retention will need to be considered in the context of the different uses of the data for research as described under the DPA 2018 for both law enforcement and non-law enforcement purposes. These considerations will allow a compliant retention schedule to be drafted by the police forces responsible for the processing of personal data to ensure directly identifiable personal information is not kept for a period that is longer than necessary. The review should cover both operational data and research data as well as other various requirements set out in law. For example, if any of the data processed was concerning a court case, this may have a direct legal impact on the data retention requirements. A full retention schedule would be the proposed output to help ensure that the next stage of processing is fully compliant with the legislation.

4.6 The requirements of retention under the GDPR allow data to be retained indefinitely if necessary, for the purpose of research. There is no such provision under part 3 of the DPA 2018, meaning that data stored for research purposes for the purposes of law enforcement must comply with the requirements of the fifth data protection principle³. This means that any data held even for research purposes under part 3 of the DPA 2018 must not be stored any longer than is necessary and should be subject to reviews at appropriate intervals for the ongoing retention of any data held.

Data sharing with police forces

4.7 During the course of the Sandbox participation, the ICO agreed with Tonic Analytics to help support compliant data sharing between the stakeholders involved in the Galileo Programme and review their data sharing agreements. Support was

³ DPA 2018 Part 3 Section 39

- (1) The fifth data protection principle is that personal data processed for any of the law enforcement purposes must be kept for no longer than is necessary for the purpose for which it is processed.
- (2) Appropriate time limits must be established for the periodic review of the need for the continued storage of personal data for any of the law enforcement purposes.

provided in respect of the creation of a data sharing agreement and how effective data sharing can take place, taking into account the ICO's current guidance on Data Sharing.

- 4.8 Prior to the ICO Data Sharing Code of Practice being published, in March 2020, Tonic Analytics provided various pieces of data sharing documentation for review to the ICO. These documents included draft and template data sharing agreements between different police forces, and between local authorities and police forces, to allow for the appropriate use of data including specific roles and responsibilities as to the provision of the rights of the data subject and any applicable determinations as to the review purpose compatibility and the original legal basis for processing where data had been collected under the scope of the GDPR.
- 4.9 Tonic Analytics requested a steer from the ICO which sought to clarify if the agreements between police forces, could enable data to be passed directly to Tonic Analytics by the data sharing police forces. The view taken by the ICO was that as long as suitable legal agreements had been put in place which include appropriate legal and technical safeguards to ensure the confidentiality of the data in transit, that it was likely still in compliance with UK data protection legislation to transfer the data between the data sharing partners and directly to Tonic Analytics, acting as data processor.
- 4.10 Additional clarity was sought by the ICO on the contractual wording used in the documents to ensure that there was an appropriate distinction made between data collected for purposes outside of law enforcement and data which has been collected for law enforcement purposes. Examples of the ICO's rationale for seeking such clarification include the fact that very early versions of the documents referred to the Article 6 lawful basis for the processing of personal data as being applicable, regardless of whether the data had been collected under the scope of the GDPR or for law enforcement purposes. The ICO also noted the importance of other language used to distinguish between special category data under the GDPR and sensitive processing under Part 3 of the DPA 2018, to prevent confusion as to the legal requirements and which part of the legislation the data controller is seeking to operate under.
- 4.11 The ICO has recently published the [Data Sharing Code of Practice](#) and further [guidance for those wishing to share data with law enforcement authorities](#).

Applicability of the exemptions from the facilitation of data subject rights

4.12 In Tonic's application to the Sandbox, they described the Galileo project as operating under the GDPR / DPA 2018 Part 2 research exemption⁴ and initially sought to rely on the accompanying exemption regarding data subject's rights (including rights of access), from the scope of the pilot stages. However, as the ICO and Tonic worked together it became clear that the processing should be considered to be taking place under Part 3 of the DPA 2018 for the extent of the processing of personal data under the research pilot stage. The exemptions referred to in GDPR/Part 2 of DPA 2018 could therefore no longer be applied and nevertheless should only be applied on a case by case basis. It is of the utmost importance that the rights of the data subject can be enforced through the relevant controller, in this case Gloucestershire Constabulary, who may choose to uphold rights requests as per the legal requirements of Part 3 of the DPA 2018.

4.13 As part of this work, the ICO advised Tonic to conduct a review of the fair processing notices provided by all stakeholder

⁴ GDPR Article 89 (2) 2. Where personal data are processed for scientific or historical research purposes or statistical purposes, Union or Member State law may provide for derogations from the rights referred to in Articles 15, 16, 18 and 21 subject to the conditions and safeguards referred to in paragraph 1 of this Article in so far as such rights are likely to render impossible or seriously impair the achievement of the specific purposes, and such derogations are necessary for the fulfilment of those purposes.

organisations associated with the Galileo Project, in relation to any processing or sharing of relevant data, in order to comply with Article 5(1) of the GDPR where the data had been originally collected outside of law enforcement purposes⁵.

- 4.14 The ICO advised Tonic that the previously available fair processing notice, provided by Gloucestershire Constabulary, may not have sufficiently informed the data subject of the activities carried out by the programme under the research pilot. Work was undertaken by Gloucestershire Constabulary, with the support of Tonic Analytics and the ICO, to revise the notices to ensure that appropriate information was publicly available going forward.
- 4.15 The ICO suggested that it may be appropriate to ensure that the aims and objectives of the Programme are well publicised to help ensure the fair processing of information and that the data subjects are sufficiently informed and can appropriately exercise their rights. In the event of the Programme moving to an operational model, the data subjects should be informed

⁵ Personal data shall be processed lawfully, fairly and in a transparent manner in relation to the data subject ('lawfulness, fairness and transparency').

specifically to the extent that is described under part 3 paragraph 44(2) of the DPA 2018⁶ with consideration of any relevant restriction of this information as is appropriate and in line with the law.

- 4.16 As the Programme moves towards operational functionality the parties should work to ensure that data subjects rights are properly upheld by the relevant authority and that agreements are in place to ensure that responsibility for the data is properly understood and documented between controllers and relevant processors. This may in particular take the form of a joint controllership agreement with references to predetermined processes to help support the facilitation of the rights of the data subject.

⁶ (2) The controller must also, in specific cases for the purpose of enabling the exercise of a data subject's rights under this Part, give the data subject the following—

- (a) information about the legal basis for the processing;
- (b) information about the period for which the personal data will be stored or, where that is not possible, about the criteria used to determine that period;
- (c) where applicable, information about the categories of recipients of the personal data (including recipients in third countries or international organisations);
- (d) such further information as is necessary to enable the exercise of the data subject's rights under this Part.

Lawfulness of the processing

- 4.17 During the course of the Sandbox participation, Tonic Analytics has consistently demonstrated a willingness to consider the effectiveness of the proposed solution, undertaking significant periods of testing with data in a pseudonymised format to demonstrate possible activities before committing to the operational processing of data.
- 4.18 The lawful basis for processing is required to be fully documented in order to comply with UK data protection legislation. After concluding that the processing was taking place under the part 3 of the DPA 2018 and under the data controllership of Gloucestershire Constabulary, it has been necessary to ensure that appropriate decision making has rested with the police as the controller, particularly as it pertains to the legal purpose for processing.
- 4.19 The current described lawful purpose for the processing pertains to the use of data for statistical purposes as outlined under part 3 of the DPA 2018, at paragraph 41⁷. The processing has been assessed by Galileo as not seeking to identify individuals

⁷ (1) This section applies in relation to the processing of personal data for a law enforcement purpose where the processing is necessary—
(a) for archiving purposes in the public interest,
(b) for scientific or historical research purposes, or

and is unlikely to cause harm or distress and is subject to other appropriate safeguards. The processing therefore meets the further requirements of processing for statistical or research purposes. Significantly, this prohibits the data processing

- (c) for statistical purposes.
- (2) The processing is not permitted if—
 - (a) it is carried out for the purposes of, or in connection with, measures or decisions with respect to a particular data subject, or
 - (b) it is likely to cause substantial damage or substantial distress to a data subject.

conducted during the course of the Sandbox participation from being used for specific operational purposes to investigate or prosecute crime as this could be considered to be a measure or decision with respect to a data subject.

- 4.20 During the course of their participation in the ICO Sandbox, Tonic Analytics also provided feedback in response to the ICO's recently published toolkit, aimed at [supporting the proper use of Data Analytics](#). This toolkit may help support the Programme as it moves to operational functionality, using identifiable personal data in a format that is usable by police to identify an individual vehicle keeper.
- 4.21 Future uses of this analytics platform to use data in an identified format to support the work of road policing will need to have a clear basis in law and operate in a fashion that is foreseeable to the data subject. The Galileo Programme is still at this time working on prototyping operational functions. It is significant that each piece of operational functionality must be assessed to ensure its legality and remains under the effective controllership of the police forces. The programme must continue to assess the necessity of the use of data analytics to achieve any legitimate and lawful aim which is the concern of the processing. This means that the processing must be the least invasive, effective and reasonable method to achieve the aim of the processing.
- 4.22 Tonic Analytics' consistent attempts to ensure that it can evidence the effectiveness of the processing in the least invasive way possible has allowed it to properly demonstrate the links between areas of non-compliance by road users with the law and the overall effectiveness of the platform, the likelihood that it may lead to better roads policing if taken forward in an operational format and the risks associated with the processing and technology.
- 4.23 Once the necessity of the processing at an operational level can be demonstrated, the Programme should look at ways in which the personal data processing may directly inform police action as a data analytics product. The Galileo Programme has outlined that once moving to an operational model, it will continue to test functionality in a fashion with a limited scope to ensure that it is continuing to look at analytics through a data protection by design lens. They should also limit the use of data to exercises where the processing of personal data is most likely to be effective and occurs in a controlled way, whilst avoiding overextending the scope without consideration as to any negative effects to the public.

- 4.24 It must also be assured by Tonic Analytics that criminal enforcement data is not being utilised in any way outside of the law enforcement purposes or passed back to any authority not considered to act as a competent authority for the basis of law enforcement processing, unless authorised by law.

5. Ending statement

- 5.1 Tonic Analytics' participation in the ICO's Regulatory Sandbox has given the ICO the opportunity to gain a valuable insight into the transportation and law enforcement sectors and how the innovative application of technology known as Digital Twin, underpinned by effective data sharing, can effectively help to tackle the challenges of reducing crime and increasing safety on the roads. The ICO's work with Tonic Analytics has increased our understanding of large-scale data sharing and the use of effective approaches for data minimisation and pseudonymisation to build a blended dataset to deliver research outcomes for the public benefit.
- 5.2 It is clear to us from our work with Tonic Analytics that they are committed to making use of innovative technology in a compliant way to provide new insights and actionable intelligence to improve the way that law enforcement organisations, public sector agencies, local authorities and the private sector can collaboratively tackle crime and safety challenges on the roads in a compliant and secure manner while maintaining individual' rights to data protection and privacy.
- 5.3 Tonic Analytics' engagement in the Sandbox included its role within the Galileo Programme, which is jointly sponsored by Highways England and the National Police Chiefs' Council, acting through its National Roads Policing, Operations, Intelligence and Investigations portfolio. In this context, the ICO's work with Tonic Analytics provided effective guidance on the use of blended data for law enforcement research purposes, including:
- The appropriate selection of the lawful basis bases for processing. In this case supporting the programme to transition from an original plan to use the DPA Part 2 Research Exemption to the more appropriate use of DPA Part 3, processing for law enforcement purposes.

- Establishing a compliant data sharing organisational structure across the multiple organisations engaged in data sharing, decision making and processing of data, with the roles of data controller and data processor being clearly identified.
- Early engagement in the ICO's Data Sharing Code of Practice to ensure appropriate data sharing agreements were in place between all parties sharing data for the Digital Twin research environment.
- DPIAs being completed in advance of any data processing or sharing activity taking place and ensuring that such DPIAs effectively identify risks and mitigating actions.
- Addressing considerations for the use for data analytics in law enforcement through use of the ICO's [Toolkit for organisations considering using data analytics](#).

5.4 Furthermore, based on the information we have seen in the Sandbox and solely in respect of the Tonic Analytics technology platform as considered in the Sandbox, it appears likely that the data processed by Tonic Analytics' technology for the purposes of law enforcement research as conducted within the Sandbox is processed securely and not in a way which breaches UK data protection legislation. Moving forward, with a perspective on continued research and the development and deployment of operational services, Tonic Analytics should ensure that they continue to follow the steers provided to them in the Sandbox, as well as relevant ICO guidance.

5.5 At the time of writing this report, the Covid-19 public health emergency was continuing to unfold across the world with unprecedented impact on the transportation and law enforcement sectors. We are immensely grateful to Tonic Analytics for their engagement in the Sandbox and finalising this report in difficult circumstances.