UNCLASSIFIED PUBLIC

VISUAL ANALYTICS FOR SENSE-MAKING
IN CRIMINAL INTELLIGENCE ANALYSIS

***VALCRI* WHITE PAPER SERIES**
VALCRI-WP-2017-010
13 February 2017
Edited by B.L. William Wong

Roadmap for the Operationalization of Legal and Privacy Requirements in VALCRI Analysis

Thomas Marquenie, and Fanny Coudert

KU Leuven Centre for IT & IP Law
Sint-Michielsstraat 6
Box 3443, 3000 Leuven
BELGIUM

Project Coordinator

Middlesex University London
The Burroughs, Hendon
London NW4 4BT
United Kingdom.

Professor B.L. William Wong
Head, Interaction Design Centre
Faculty of Science and Technology
Email: w.wong@mdx.ac.uk

Copyright © 2016 The Authors and Project VALCRI. All rights reserved.

INTENTIONALLY BLANK

ABSTRACT

This White Paper seeks to illustrate how the LEP guidelines and several legal requirements are set to be further operationalized in the VALCRI project. Based on VALCRI's data management policy for the live data tests of February 2017, it formulates technical solutions and concrete measures to meet legal data protection requirements adapted to the working environment of criminal analysts.

Keywords

Visual Analytics, Criminal Intelligence Analysis LEP Guidelines, Crime, Analytics, Data Management, Data Protection Principles, Operationalization.

INTENTIONALLY BLANK

INTRODUCTION

With the development of new technologies and the expansion of available datasets, law enforcement agencies (LEAs) are increasingly reliant on Big Data analytics and intelligence-led policing for the execution of their tasks. It is possible to improve the efficiency of policing activities such as criminal intelligence analysis by recording and making use of large amounts of data. Generally speaking, criminal analysis services of the police process personal data for three different purposes, being strategic, tactical and operational analysis [1]. Strategic analysis seeks to support and develop general law enforcement policies contributing to the efficient and effective prevention of, and the fight against, crime. It delivers general information on criminal statistics and trends, hereby providing a high-level pictures of given crime phenomena. Tactical analysis aims to manage and coordinate day-to-day police tasks and assist police personnel in the execution thereof. Among others, it supports the identification of priority targets such as geographical areas and type of crimes. Operational analysis takes place within the context of specific criminal investigations and seeks to assess and establish ties between criminal acts, suspects, *modi operandi* and other elements related to the investigation and prosecution of a crime.

To this end, the VALCRI project (Visual Analytics for Sense-making in Criminal Intelligence Analysis) seeks to develop a system to assist in the information management and data evaluation conducted by criminal intelligence analysts in law enforcement agencies [2]. VALCRI will serve as a working space which extracts meaningful information from text, documents, images and video, detects patterns and presents such data in an accessible and insightful way to its user. By coupling visualization and computation, VALCRI allows analysts to draw conclusions and make connections otherwise often missed by humans.

As intelligence-based systems process large amounts of information capable of affecting the lives and rights of individuals, including the fundamental right to privacy, therefore the risks they create should be addressed and properly mitigated [3]. For this purpose, the VALCRI project has set up a working group dedicated to the **S**ecurity, **E**thical, **P**rivacy and **L**egal ('SEPL') aspects of the system. In order to meet practical, legal and ethical requirements, the SEPL group drafted general guidelines in order to support VALCRI's further development [4]. On the basis of the applicable legal framework, these guidelines identified 34 legal, ethical and privacy-related (LEP) requirements for developers to incorporate in the design of VALCRI. These guidelines, an overview of which can be found in appendix A, have been updated, wherever necessary, to take into account the provisions of the recently adopted Police and Justice Data Protection Directive (Directive (EU) 2016/680) which will become the new EU data protection standard for law enforcement data processing operations in 2018. The guidelines also provide a brief description of the issues or related legal obligations they intend to address and give a possible method to adequately address them.

The LEP guidelines have then been operationalized on the basis of the specific working environment of criminal analysts. To that end, the end-user partners to the project, being the Local Antwerp Police and West Midlands Police Force set to adopt the VALCRI system upon its completion and participate in its preliminary tests, have been engaged in a series of face-to-face meetings. The goal of these discussions has been twofold: 1) to understand how the requirements are currently implemented in these two police environments and 2) to identify how these guidelines need to be refined and translated into technical requirements adjusted to the new technical environment provided by VALCRI. This second phase resulted in the drafting of a Data Management Policy for the first tests of VALCRI prototype.

This White Paper presents the outcome of this process. It is intended to provide VALCRI developers with concrete solutions on how to deal with abstract and general legal requirements for the design of the system. To that end, it draws on the LEP guidelines, the relevant national and European legislation, and the data management policy. As such, this document should be considered a basic roadmap presenting how general legal, ethical and privacy-related requirements are to be converted into concrete and directly applicable measures for the VALCRI system. It shall soon be supplemented by a document on the current status of the implementation of these requirements. Finally, it should be noted that this White Paper does not refer explicitly to the applicable legal provisions, as they are already an integral part of the LEP guidelines.

The following section describes each relevant legal principle, a discussion of the current practices in policing, guidelines developed in VALCRI for their interpretation and use, and the requirements they present when we implement them in VALCRI. These principles are:

1. Purpose limitation and data minimization
2. Categories of data subjects
3. Traceability and accountability
4. Data accuracy

5. Anonymization and pseudonymization
6. Data storage and deletion
7. Security measures

Appendix B tabulates and summarises the details in this section for easy reviewing.

REQUIREMENTS AND TECHNICAL FEATURES

Principle 1: Purpose limitation and data minimization

Definition. A key data protection principle is that of purpose limitation. Under this principle, the purposes for which data can be processed must be determined prior to its collection. The original purposes shall serve as a binding guideline requiring all future processing activities to be compatible with these objectives [5]. Additionally, following the principle of data minimization, the data must be adequate, relevant and limited to what is necessary for the purposes for which it was collected and further processed. This purpose must be specified, made explicit and be legitimate in nature. The proper documentation of the purpose is both an autonomous requirement in its own right as well as a necessary tool serving the assessment of the requirements of legitimacy and transparency. As such, the data controller must be sufficiently precise when determining the purposes which are to be established in a clear and intelligible way before the start of the processing activities.

Guidelines. Guideline G3.4.P5 recognizes the necessity of the enforcement of purpose limitation. The VALCRI system must be able to support the limitations put in place by police services when defining the purposes of the data collection and processing and establishing the competences and tasks of their criminal analysts on the basis of the applicable legal framework, being national laws on the functioning of police, data protection and criminal procedure.

Current Practices in Policing. Law enforcement agencies typically set out a number of legitimate tasks and objectives for which personal data may be processed in a lawful manner [6]. These tasks serve as a legal basis and legitimate purpose for the processing activities. In light of these tasks, the data necessary to fulfill them is then determined prior to the processing and tied to the profile of the analyst. User profiles will thus be dependent on the tasks assigned and the level of clearance. Currently, once these legitimate processing purposes are identified, there is no recurring obligation to continuously document the purpose for each particular processing action, meaning that police analysts are permitted to process personal data for their daily tasks. Thus, adherence to the requirement of purpose limitation will mainly require the implementation of a strict access control policy and the full traceability of the actions performed by the analyst.

Implementation in VALCRI. At the top level, a login shall be required for users to access the system. By signing into the system, all actions performed by the analyst are logged and tied to his or her user profile which ensures accountability and allows compliance with the principle of purpose limitation to be established. These logs shall not be freely accessible and shall be auditable only under specific conditions and by the competent oversight authority. In addition, the role of the user in the law enforcement authority and the nature of a particular case can be used to further limit what data a certain user can access. Different user levels with differing degrees of access to certain data might be created in order to differentiate between supervisory authorities, administrative personnel, researchers, system administrators and several types of criminal analysts, as far as they are distinguishable within the police force. The access level of the user and his or her tasks and obligations serve as a sufficient purpose for the processing activities. They allow the analyst to process the relevant information without requiring the a priori designation of a new purpose for each processing act. Additionally, in order to improve traceability without impeding on the work of the analysts, VALCRI should give the possibility to link search queries to a specific and predefined tasks.

Principle 2: Categories of data subjects

Definition. Criminal analysis requires the processing of different categories of persons with different degrees of involvement in the crime. The processing of personal data in the context of criminal intelligence analysis must take into account the role of the data subject in the criminal act. In practice, this means that personal data relating to a victim or a witness of a crime cannot be treated as if it concerned a suspect or known convict. As their relation to the criminal fact is fundamentally different and warrants a different treatment of their data, it is necessary that such a distinction is maintained in practice. The qualification as suspect, contact or witness of an individual results in different consequences for his or her fundamental rights [7]. Errors in the qualification might have a serious impact on the individuals concerned [8].

Guidelines. Guideline G3.4.P8 recognizes the need for a distinction between data subjects based on their role in the investigation or their relation to the crime, as the seriousness of the interference in their private life should correlate with their role in the investigation. As required by the European

Police and Criminal Justice Authorities Directive (Directive (EU) 2016/680), such a distinction must be made between:

(a) persons with regard to whom there are serious grounds for believing that they have committed or are about to commit a criminal offence;

(b) persons convicted of a criminal offence;

(c) victims of a criminal offence or persons with regard to whom certain facts give rise to reasons for believing that he or she could be the victim of a criminal offence, and

(d) other parties to a criminal offence.

Implementation in VALCRI. In VALCRI, this shall be accomplished by maintaining a clear distinction between certain categories of data subjects based on the nature of their involvement with the criminal fact. As such, this legal requirement can be met through the implementation of a tagging system which allows persons and the personal data relating to them to be marked, sorted and categorized based on their relation to the criminal fact as described above.

Principle 3: Traceability and accountability

Definition. The traceability of data in a system and the ways in which it has been processed is an important part of adequately protecting personal data [9]. Not only does it aid the users of the system in the efficient exercise of their tasks, but it allows the misuse of the system and data therein to be monitored and addressed in order to ensure accountability of both the system and its users. For a system to be accountable, it must be able to be demonstrated that it does not behave in a non-permissible or faulty manner. For the user, he or she must be able to account for and justify certain actions made or decisions taken. Additionally, a high degree of traceability assists in maintaining and validating the accuracy and authenticity of the data by monitoring it for unwarranted modifications. For the purpose of improving accountability and data accuracy by guaranteeing a sufficient manner of traceability, the Police and Criminal Justice Directive therefore implements a strict requirement of logging practices.

Guidelines. The issue of traceability and accountability is covered by a number of LEP guidelines. Among others, G3.4.PL3 on auditing and logging, G3.4.L2 on accountability and G3.4.LEP1 on system transparency make direct reference to how the implementation of certain policies can improve the traceability of data in VALCRI as well as support accountability of both the user and system.

Implementation in VALCRI. In VALCRI, sufficient traceability and accountability shall be achieved through the implementation of extensive logging measures. This will be implemented through a process recording and keeping detailed logs of all relevant user and system actions in order to support the three dimensions of provenance present in VALCRI, being data, process and reasoning provenance [10]. To be fully effective, the practice of logging is comprised of two elements. First, logs shall be kept of the actions of the system, as the user and system administrator must be able to determine why the software executed a certain action or how it arrived at a specific conclusion. Second, user interactions in the system shall be registered as well. The searches executed, data accessed and information processed by a certain user must be logged and tied to the identifiable person. Among the necessary information to be recorded by VALCRI are the identity of the user, time and date of the system being accessed, data gathered and accessed by the user, modifications and additions made to the analyst files, transactions of data, changes in user settings, and the processing and reasoning actions of the system. Due to the broad scope and extensive nature of these logs, there is no need to implement separate logging practices and measures for different types of analyst. Yet, in the context of access to the content of the logs, different access practices can be established in order to distinguish between, for example, technical personnel and oversight authorities. While it is key to maintain the distinction between different access profiles and link actions and operations to a specific user, the scope and content of the logs kept shall be the same for strategic, tactical and operational analysts. This functionality shall allow both analysts to retrace steps taken by the system, without being able to modify these logs, and supervisory authorities to examine how the system behaves in order to guarantee reliability, accuracy and authenticity.

Principle 4: Data accuracy

Definition. Another key principle of the protection of personal data is that of data accuracy. For such data to be processed lawfully, it must be accurate and, wherever necessary, kept up to date. In principle, only data which is reliable and accurate can be processed in a lawful and legitimate manner [11]. Personal data which is found to be inaccurate, outdated or incomplete in light of the purposes for which it is collected and processed must therefore be corrected or erased without delay.

Guidelines. Guideline G3.4.P9 recognizes the importance of maintaining a high level of data accuracy in the VALCRI system. As a general guideline, it suggests that the presumed accuracy of information is attached to the information itself and that the system, as far as possible, only processes accurate and reliable data. Yet, the guideline acknowledges that only processing verified and entirely reliable information is not always a realistic approach for criminal investigations. Because of this, other safeguards are to be put in place to protect the rights of the data subject and the integrity of the investigation. In order to further support the accuracy of personal data, guideline G3.4.P7 establishes the need of data changes in source databases being updated and synchronized in the VALCRI system. As such, alterations made in police datasets shall be accurately reflected in the system.

Implementation in VALCRI. To ensure that the data processed by VALCRI is sufficiently accurate, a number of technical measures must be implemented.

First, data must be marked appropriately based on its reliability. A distinction must be maintained between different types of data which shall subsequently be marked accordingly. Data representing proven facts, hypotheses, intelligence and witness statements shall be designated as such and differentiated between in order to avoid uncertainty and allow the analyst to assess the quality, reliability and accuracy of the data. To this end, VALCRI must be able to incorporate the pre-existing 4x4x4 (Belgium) and 5x5x5 (UK) grid structures used to assess the reliability and quality of criminal intelligence. While this currently poses certain issues for source databases used by police forces which do not utilize such a classification similar, these problems are expected to be solved as law enforcement agencies across the EU adjust to the new requirements of the Police and Criminal Justice Authorities Directive by 2018. As such, VALCRI shall be adequately equipped to meet the conditions of this new Directive at the time it enters into force.

Second, VALCRI must regularly synchronize with the source databases to detect and reflect changes occurring therein. The access to and management of police databases is strictly regulated and subject to specific procedures. Data present in such databases or police archives must be supplied in a particular way and cannot be directly altered by persons retrieving it for the purpose of police investigations. VALCRI shall never be able to change data in the source database, but shall instead react to changes made therein and present them to the user of the system. As such, in the event that data is altered, deleted, merged or updated in the source database, whether as result from further police investigations or a data subject exercising his or her rights, the system shall be able to notify the user of these changes, inform him or her of the modifications made and display the updated information.

Third, the system shall allow users to make notes in their work files within the VALCRI User Interface relating to the accuracy of data they encounter. While they cannot directly alter data in the source database, the system shall be able to create a temporary entity reflecting the updated information while the analyst relies on other procedures to contact the owner of the source database about the reliability of the data therein.

Principle 5: Anonymization and pseudonymization

Definition. In order to fully protect the personal data of persons whose information is being processed, it is required that personal data is only kept in a form allowing for the identification of the data subjects for as long as necessary for the purpose for which the data were collected [12]. As such, measures must be taken to avoid the limitless storage of personal data which is no longer relevant to the purpose of the processing activities. Anonymization and pseudonymization techniques have been identified as ideal solutions to meet this legal requirement. Data is considered successfully anonymized when it is stripped of the elements capable of identifying the natural person and the re-identification thereof is entirely impossible, at which point it can no longer be considered personal data [13]. Contrary to this, pseudonymization only conceals the identity of the data subject in a retraceable way and allows for later re-identification. While pseudonymized data is still considered personal data, its processing is considered a less significant interference with the data subject's private life.

Guidelines. Guidelines G3.4.P2 and G3.4.P3 intend for personal data to be anonymized as much as possible to correspond with the principles of necessity and data minimization. If full anonymization would not be an option, data must be pseudonymized whenever possible to further limit the interference in the private sphere.

Current Practices in Policing. The degree to which data should be anonymized depends on the tasks and status of the particular analyst. Anonymized data shall suffice and therefore be required for strategic and, under usual circumstances, tactical analysisists whose tasks include the drafting of statistical analyses or detecting crime patterns. On the other hand, the full personal data shall be necessary for the tasks of operationalist analysts who work on a specific file. For the latter, pseudonymization should be applied whenever possible and data should be anonymized for further storage after the legal retention period expires.

Implementation in VALCRI. In the VALCRI system, anonymization of data is set to be achieved through two separate methods in order to meet the different preferences and needs of police personnel. First, the system shall be able to only present anonymized data to the user by means of specifically structured queries. By employing specific queries relating to a particular task, the user can request the system to only retrieve the necessary data for the purpose of a certain processing activity. If the query is structured in such a way that the user's task does not require the full personal details of persons involved, the system shall present anonymized or pseudonymized data. Second, VALCRI shall contain a number of PET (Privacy-Enhancing Technology) functionalities which allow for further anonymization and pseudonymization of the personal data contained in the system when the user chooses to apply them as such [14]. These functionalities shall include, for example, face detection and obfuscation techniques applicable to video data.

Principle 6: Data storage and deletion

Definition. Serving as an extension to the general principles of purpose limitation and data minimization further described above, personal data cannot be stored longer than strictly necessary for achieving the initial purpose of collection. As such, data must be deleted whenever it is no longer relevant for this purpose. Procedures must be in place to support the timely assessment and deletion of data which is either no longer consider relevant and necessary or has been stored for the maximum period allowed for by law. This period of retention may vary based on the crime, type of database, police force processing the information and the purpose of the processing. In the event that data is no longer considered necessary, it can be stored when fully anonymized.

Guidelines. Guideline G3.4.P6 recognizes the need for the timely deletion of data. It suggests that a specific data retention concept must be developed for types of personal data and that existing rules on the deletion of data must be implemented into VALCRI so that the system can play a supporting role therein. According to the guideline, data deletion must take place after the fulfilment of the purpose for which the data was collected or according to the legally binding storage periods.

Implementation in VALCRI. For VALCRI, a distinction must be made between two instances of data deletion. First, the data stored in other police databases which are retrieved and presented by the system, cannot be modified or deleted by the system. VALCRI must incorporate the functionalities described under data accuracy in order to notify the user of erasures taking place in the databases from which VALCRI draws its data.

Second, the personal data stored by the system in analyst work files tied to a particular case should be subject to the same duty of timely deletion. The system shall introduce semi-automated procedures notifying the users when data stored in the system should be removed. As such, VALCRI must be configured in accordance with the national laws and data retention procedures, ensuring that data is not kept longer than necessary or allowed for by law. Following this, VALCRI shall notify users the data which is due for erasure. Typically, the data stored in the work files belonging to operational analysts are stored externally even after the completion of the analysis report. Only while the analyst is working on a specific case is this data stored in the VALCRI system itself. When the investigation is declared closed, the information and correlating log data are stored and archived for the legal duration before its final erasure. These closed files are not archived in VALCRI but are instead stored in specific archives with limited access.

Principle 7: Security measures

Definition. Finally, police information systems must contain sufficient safeguards and adequate security measures in order to protect and guarantee the integrity, safety and security of the police investigations and the personal data processed in the context thereof. To protect the system against loss and destruction of the data in the system, as well as the unauthorized or unlawful processing thereof, technical and organizational measures must be taken which take into account the technological state of the art, the costs of implementation and the nature, scope and purposes of the processing activities.

Guidelines and implementation in VALCRI. In VALCRI, a number of different security goals have been identified. In order to ensure that the VALCRI system and the data therein are highly secure, the so-called information security objectives of confidentiality, integrity, availability, accountability, assurance and non-repudiation serve as a fundamental guideline through the development process [15].

As such, a security policy consists of several separate measures, a number of them have already been discussed above and can be found both in different LEP guidelines and earlier VALCRI deliverables [16]. Measures such as adequate access control (AC) to limit access ex-ante to the structured and unstructured data across system domains (G3.4.PL4), data synchronization (G3.4.P7), detailed high-assurance and analytical logging and auditing systems recording both user and system actions (G3.4.PL3), and strict procedures for the deletion and modification of data in VALCRI (G3.4.P6) shall contribute to a comprehensive security policy and must be implemented in VALCRI. Additionally, adequate measures protecting the system during all stages of the processing activities must be in place. Privacy-enhancing technologies shall be implemented to secure the pre-processing and modification of certain personal data in the system. Data-in-transit shall be protected during transmission by means of encryption. Back-up and recovery functionalities shall be provided for and used to their full extent, while the VALCRI system itself and all transmissions of data thereto or therefrom shall remain secure. System malfunctions shall not be able to corrupt the integrity of the data and the actions of the system shall be able to be evaluated and assessed on their reliability to ensure that VALCRI is working properly and as intended.

Nevertheless, technical measures can only accomplish so much to improve the security of a system, as the human component is another key aspect of data protection and security. As such, organizational measures are equally important and necessary. Physical access control, data management policies, safeguards for the transport and management of data media and equipment, and adequate training of police personnel shall all be the responsibility of police services using the VALCRI system.

REFERENCES

[1] VALCRI Deliverable D2.2 - Requirements Analysis.

[2] VALCRI FP7 Proposal.

[3] VALCRI Deliverable D3.4 - Human Issues Framework.

[4] VALCRI Legal, Ethical & Privacy Guidelines, M16.

[5] Working Party 29 Opinion 03/2013 on Purpose Limitation, 2013.

[6] VALCRI Data Management Policy Belgium.

[7] Working Party 29 Opinion 03/2013 providing further input into the discussions on the draft Police and Criminal Justice Data Protection Directive, 2013.

[8] ECtHR, *Dimitrov-Kazakov v. Bulgaria*, 10 February 2011.

[9] J. ALHADEFF, B. VAN ALSENOY and J. DUMORTIER, “The Accountability Principle in Data Protection Regulation: Origin, Development and Future Directions” in D. GUAGNIN et al., *Managing Privacy through Accountability”*, 2012.

[10] VALCRI White Paper Analytical Provenance for Criminal Intelligence Analysis.

[11] L. CAI and Y. ZHU, “The challenges of data quality and data quality assessment in the Big Data era”, *Data Science Journal*, 2014.

[12] Information Commissioner’s Office, *Anonymisation: managing data protection risk – code of practice*, 2012.

[13] Working Party 29 Opinion 05/2014 on Anonymisation Techniques, 2014.

[14] VALCRI Deliverable D11.14 – Data Analysis: Synthetic Data Creation and Management Report.

[15] VALCRI Deliverable D2.3 – System Design (v1).

[16] VALCRI Deliverable D12.18 – System Integration: S2DP Architecture and Implementation.

ACKNOWLEDGEMENTS

We wish to thank our police partners for allowing us unfettered access to their staff and time, and for helping us answer many difficult legal and technology implementation questions.

APPENDIX A – LEP GUIDELINES

A detailed discussion of each of these 34 principles can be found in the deliverable D3.4 Human Issues Framework.

LEGAL

Guideline ID	Title
G3.4.L1	Respect for fundamental rights and freedoms
G3.4.L2	Accountability
G3.4.L3	Intellectual Property (e.g. Copyright)
G3.4.L4	Function creep
G3.4.L5	Territorial jurisdiction
G3.4.L6	Administrative safeguards
G3.4.L7	Electronic signatures
G3.4.L8	Enabling correct documentation of cases
G3.4.L9	Inspection report
G3.4.L10	Management of data with risks attached

ETHICAL

Guideline ID	Title
G3.4.E1	Human dignity and autonomy DETECTOR project – Data Mining (profiling) example.

PRIVACY

Guideline ID	Title
G3.4.P1	Personal information
G3.4.P2	Anonymization of data
G3.4.P3	Pseudonymisation of data
G3.4.P4	Limited use of sensitive personal data
G3.4.P5	Enforcement of purpose limitation
G3.4.P6	Data deletion
G3.4.P7	Synchronisation of data changes
G3.4.P8	Differentiation of the roles of the data subjects
G3.4.P9	Accuracy of data
G3.4.P10	Data for administrative purposes
G3.4.P11	Effective exercise of data subjects' rights
G3.4.P12	Privacy of employees
G3.4.P13	Restricted access of system administrators
G3.4.P14	Avoidance of free-text fields
G3.4.P15	Prohibition of fully automated final decision-making

PRIVACY & LEGAL

Guideline ID	Title
G3.4.PL1	Legal Compliance
G3.4.PL2	Data Limitation: Necessity principle
G3.4.PL3	Auditing and logging
G3.4.PL4	Access control

LEGAL and ETHICAL

Guideline ID	Title
G3.4.EL1	Equality: No prejudice
G3.4.EL2	Truthfulness: Verifiability and counter-examination

Remaining and overlapping LEGAL, ETHICAL & PRIVACY

Guideline ID	Title
G3.4.LEP1	Transparency of the system
G3.4.LEP2	Chain of custody and provenance of data

APPENDIX B – SUMMARY OF LEP GUIDELINE IMPLEMENTATION

Principle	LEP Guidelines	Requirements	Remarks
1. Purpose limitation and data minimization: The purposes for which data can be processed must be determined prior to its collection. All future processing activities which must fall within the scope of the original purpose. The data must be adequate, relevant and limited to what is necessary in relation to these purposes for which it is collected or further processed.	G3.4.P5: Enforcement of purpose limitation	Unique user profiles which are tied to individual analysts and allow for varying degrees of access to data based on the tasks assigned and clearance given. Detailed access control measures, extensive logging procedures allowing for audits, and the possibility of linking search queries and accessible data to predefined tasks and access levels should be provided.	
2. Categories of data subjects: The processing of personal data in the context of criminal investigations must take into account the role of the data subject in the criminal act and make distinctions based on the nature of his or her involvement.	G3.4.P8: Differentiation of the roles of the data subjects	Maintaining a clear distinction between certain categories of data subjects based on the nature of their involvement with the criminal fact by implementing a tagging system which allows persons and the personal data relating to them to be marked, sorted and categorized based on their relation to the criminal fact as described above.	
3. Traceability and accountability: Both system and user actions must be recorded and traceable for auditing purposes. System errors and unlawful use must be identifiable and determining accountability must be possible in cases of misuse in order to guarantee reliability, accuracy and authenticity.	G3.4.L2: Accountability G3.4.PL3: Auditing and logging G3.4.LEP1: Transparency of the system	Implementing extensive, detailed and auditable logs of all system and user actions. These logs shall include the identity of the user, time and date of the system being accessed, data gathered and accessed by the user, modifications and additions made to the analyst files, transactions of data, changes in user settings, and the processing and reasoning actions of the system.	
4. Data accuracy: For data to be processed lawfully, it must be reliable, accurate and, wherever necessary, kept up to date. Data which is inaccurate, outdated or incomplete in light of the processing purposes must be corrected or erased.	G3.4.P7: Synchronization of data changes G3.4.P9: Accuracy of data	Implementing the possibility of marking data based on reliability (4x4x4 or 5x5x5 grids) and distinguishing between different types of data (facts, hypotheses, intelligence and witness statements). Synchronizing data with source databases to detect and reflect changes, as well as notify users of modifications made. Allowing users to make notes within the VALCRI UI relating to data accuracy and create temporary entities reflecting data as updated by the user.	It is not yet clear in VALCRI how this is going to be solved. Notifications within a certain predefined time interval is one possibility, flagging/blocking/marking of data another. Open issue at the moment.
5. Anonymization and pseudonymization: Personal data may only be kept in a form allowing for the identification of the data subjects for as long as necessary for the purpose for which the data were collected. When identifiable information is not necessary for a specific purpose, the relevant data must be presented in an anonymized or pseudonymized manner.	G3.4.P2: Anonymization of data G3.4.P3: Pseudonymization of data	Implementing the possibility of only presenting anonymized or pseudonymized information based on the tasks and status of the analyst, as well as the search queries used. Introducing additional PET functionalities and anonymization technologies to allow users to further anonymize specific data.	This is not yet clearly understood. To be discussed with FHG to see if this description is accurate for search queries.
6. Data storage and deletion:	G3.4.P6: Data deletion	Incorporating the measures described under data accuracy to inform users of	

Personal data cannot be stored longer than strictly necessary. Only when the data is relevant to the purposes of the processing may it be stored in information systems. Data which is either no longer consider relevant, necessary or has been stored for the maximum period allowed for by law must be deleted in a timely fashion.		deletions occurring in source databases. Implementing semi-automated procedures notifying the users when data stored in analyst work files should be removed. Allowing configuration of procedures to support the timely assessment and deletion procedures in accordance with national laws and data retention procedures to ensure this data is kept no longer than necessary or allowed for by law.	
7. Security measures: Information systems must contain sufficient safeguards and adequate security measures in order to protect and ensure the integrity, safety and security of the police investigations and the personal data processed in the context thereof.	G3.4.P6: Data deletion G3.4.P7: Synchronization of data changes G3.4.PL3: Auditing and logging G3.4.PL4: Access control	Implementing adequate security measures as described in other requirements and VALCRI deliverables. These measures shall include access control protocols to limit access to structured and unstructured data across system domains, data synchronization, logging and auditing systems, and data deletion and modification procedures. Additionally, system security, PET, encryption, back-up and recovery functionalities shall be implemented.	

The research leading to the results reported here has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) through Project VALCRI, European Commission Grant Agreement Number FP7-IP-608142, awarded to Middlesex University and partners.

	VALCRI Partners	**Country**
1	Middlesex University London Professor B.L. William Wong, Project Coordinator Professor Ifan Shepherd, Deputy Project Coordinator	United Kingdom
2	Space Applications Services NV Mr Rani Pinchuck	Belgium
3	Universitat Konstanz Professor Daniel Keim	Germany
4	Linkopings Universitet Professor Henrik Eriksson	Sweden
5	City University of London Professor Jason Dykes	United Kingdom
6	Katholieke Universiteit Leuven Professor Frank Verbruggen	Belgium
7	A E Solutions (BI) Limited Dr Rick Adderley	United Kingdom
8	Technische Universitaet Graz Professor Dietrich Albert	Austria
9	Fraunhofer-Gesellschaft Zur Foerderung Der Angewandten Forschung E.V. Mr. Patrick Aichroft	Germany
10	Technische Universitaet Wien Assoc. Prof. Margit Pohl	Austria
11	ObjectSecurity Ltd Mr Rudolf Schriener	United Kingdom
12	Unabhaengiges Landeszentrum fuer Datenschutz Dr Marit Hansen	Germany
13	i-Intelligence Mr Chris Pallaris	Switzerland
14	Exipple Studio SL Mr German Leon	Spain
15	Lokale Politie Antwerpen	Belgium
16	Belgian Federal Police	Belgium
17	West Midlands Police	United Kingdom