Towards a Linked Information Architecture for Integrated Law Enforcement

Wolfgang Mayer[1], Markus Stumptner[1], Pompeu Casanovas[2,3], and Louis de Koker[2]

[1] University of South Australia, Adelaide, Australia
[2] La Trobe Law School, La Trobe University, Melbourne, Australia
[3] UAB Institute of Law and Technology, Universitat Autònoma de Barcelona, Spain

Abstract. Law enforcement agencies are facing an ever-increasing flood of data to be acquired, stored, assessed and used. Automation and advanced data analysis capabilities are required to supersede traditional manual work processes and legacy information silos by automatically acquiring information from a range of sources, analyzing it in the context of on-going investigations, and linking it to other pieces of knowledge pertaining to the investigation. This paper outlines a modular architecture for management of linked data in the law enforcement domain and discusses legal and policy issues related to workflows and information sharing in this context.

Keywords: law enforcement, investigation management, linked data.

1 Introduction

Investigations conducted by law enforcement agencies (LEAs) are increasingly reliant on effective collection and analysis of information that may be obtained from a variety of sources, internal and external to the organization [1]. Investigations generally follow an iterative process of information collection, assessment, investigation planning, execution, and brief of evidence preparation where each step either produces new information or relies on information collected earlier in the process.

Information collected within the organization include information about individuals, organizations, objects and entities of interest, witness statements, evidence obtained from crime scenes, communications intercepts and the results of forensic analysis. This information may be complemented and integrated with data such as financial transactions, travel and immigration records, and criminal history that are obtained from external sources. In addition, documentation about the investigation process and data provenance must be maintained in order to establish that evidence submitted to court had been obtained within the law and policies relevant to the investigation.

Accessing data as well as linking and integrating them in a correct and consistent way is a pressing challenge in particular when underlying data structures and access methods change over time. Lack of interoperability between information systems

within and across organizations remains one of the prevalent concerns of investigators [2]. Investigations are delayed by poor information management practices that result in information being unavailable or not being available in a timely manner, poor information quality, and cumbersome manual approval and information retrieval procedures.

The project Integrated Law Enforcement (ILE), conducted by the Data to Decisions Cooperative Research Centre (D2D CRC)[1], aims to develop a platform where investigators can manage the information collection, analysis, and processes pertaining to a case through a consistent single user-facing platform. The project has been developing technological solutions for information management, linking, and analysis that are tailored to the needs of investigators. An extensible software architecture for searching, linking, and integration of data sources forms one of the corner stones of the project. The platform will eventually include analytic services that can be invoked by investigators. The data management architecture is complemented with a state of the art user-facing portal and an analysis of legal aspects pertaining to workflows and information sharing.

Effective linking, integration, and analysis of data requires breaking down data "silos" and opening up legacy systems within organizations to make information accessible, establishing procedures and technical infrastructure to effectively and timely share information across organizational boundaries, creating data standards to facilitate interpretation and analysis of the body of collected data, and automating, where possible, analysis and semantic enrichment of data [3].

Data integration in this context raises serious legal compliance and good governance challenges. Compliance with existing laws and principles is a pre-condition of the whole process [4]. Transparency and privacy should be preserved to foster trust between citizens and national security and law enforcement agencies. A 2015 literature review on online data mining technology intended for law enforcement broadly singled out eight main problems (crimes, investigative requirements) in 2015 [1]. Separately, some criminologists warned against the profound effect of automated data collection on the traditional criminal justice system, as it could undercut the due process safeguards built into the traditional criminal justice model [5]. It is our contention that this technological modelling should be performed under the protections of the rule of law.

In this paper, we present the overall system architecture for information sharing and outline the related legal issues pertaining to workflows and information exchange in the context of policing investigations. Our work has resulted in a data access framework for law enforcement which provides a comprehensive data and meta-data model including provenance, security, confidence, links and timeline information related to entities and links. This meta-data layer spans a Knowledge Graph-like view [7] of information pertaining to entities relevant to investigations. The resulting data and meta-data model serves as the foundation for information use, governance, data quality protocols, analytic pipelines and exploration of search results.

[1] http://www.d2dcrc.com.au/

2 Information Sharing in Law Enforcement

Timely information sharing domain is crucial for the success of many investigations in the law enforcement domain. Unfortunately, many investigations are stalled by one or more of a number of impediments related to effectively sharing information among investigators and organizations [2]. In the following we highlight a selection of issues relevant in context of linked information access.

Among the *technical impediments*, internal information silos and cumbersome information access procedures are common. Investigators routinely enter the same queries across a multitude of legacy information systems and manually collate and integrate the results. Lack of information access mechanisms for investigators in the field hamper the timely acquisition of information in electronic form and information may not be updated in timely manner. In absence of automated alerts, investigators may be unaware that new information relevant to a case has become available unless they manually issue queries periodically or rely on informal personal connections to receive notifications. As a result, relevant information may be missed even though it had been available in an information system. Data quality varies greatly as data quality standards are often not enforced and instead left to the individual user.

Workflows and policies may impact upon investigations. Where approvals for actions are required, for example expenditure approval for call records requests, antiquated policies and work processes may still rely on paper forms and manual approval which result in excessive delays, in particular if approvals are sought outside of normal office hours. Here, automation and electronic means of requesting and obtaining warrants and approvals would streamline the investigation process.

Legal issues relate to restrictions on information use and sharing. For example, information obtained under a warrant for a specific investigation may not generally be used in the context of other investigations. Similarly, agencies are generally subject to restrictions on what information they can share with other agencies [2]. Even where information sharing may be legally permitted, many organizations, concerned about the implications of breaching the law, are prone to adopt prudential attitudes and policies that perhaps may unnecessarily restrict what can be shared. *Information security* and access control are challenging issues when multiple systems and organizations are involved. It is challenging to guarantee comprehensive and secure access to a large number of users accessing a multitude of information systems across organizational boundaries. Moreover, there is interaction between analytics and security attributes as new information derived from automated analytic processes must be classified using appropriate security policies to avoid inadvertently disclosing otherwise inaccessible information. Determining appropriate classification and access restrictions can be challenging in organizations.

3 System Architecture

An open architecture for data/meta-data management and analytic processes has been defined. It translates the best practices from Enterprise Application Integration to the

"Big Data" analytic pipelines [6]. Our work addresses aspects related to data and meta-data modelling and storage, modelling and execution of analytic processes, and efficient execution of analytic processes across multiple analytic tools and data sources. Central to this architecture is a method for effective semi-interactive entity linking and querying of linked data (akin to automatically generated linked ontologies such as YAGO [13]). The project intends to realize a comprehensive data management framework that relies on a well-defined share data and meta-data model supported by vendor-agnostic interfaces for data access and execution of processes comprising analytic services offered by different tools.

The overall architecture of the ILE platform is shown in Figure 1. A federated architectural model has been adopted, where one or more instances of the ILE platform can be deployed and access a number of external data sources. Each instance may provide query and analytic services to the front-end applications and can obtain data from other instances and external sources on demand. This approach is necessary as data in external sources is usually controlled by external organizations and may change at any time. Moreover, organizational policies in this context rarely support traditional Extract-Transform-Load ingestion processes across organizational boundaries.

The ILE platform provides *programmatic interfaces (APIs)* to front-end applications to access data and invoke analytic services. The interfaces expose the platform's services using a uniform data format and communication protocol. The APIs can be accessed from a desktop front-end where investigators can search and enter information as well as invoke services. Mobile applications for investigators may be developed in future versions of the platform.

Each instance maintains a *Curated Linked Data Store*, that is, a set of databases that collectively implement a knowledge-graph like structure comprising entities and their links and meta-data. This curated data store holds facts and meta-data about entities and their links whose veracity has been confirmed. This data store is used to infer the results for queries and to synthesize requests to external sources and other instances if further information is required. As such, the linked data store implements a directory of entities and links enriched with appropriate meta-data and source information such that detailed information can be obtained from authoritative sources that may be external to the system. This approach is needed as data in the law enforcement domain is dispersed among a number of systems owned and operated by different agencies. As such no centrally controlled database can feasibly be put in place in the foreseeable future.

The information contained in the linked data store is governed by an *Ontology* that defines the entity types, link types, and associated meta-data that is available among the collective platform. The ontology acts as a reference for knowledge management/organization and aids in the integration of information stemming from external sources, where it acts as a reference for linking and translating information into a form suitable for the knowledge hub. The ontology has been designed specifically for the law enforcement domain and includes detailed provenance information and meta-data related to information access restrictions. It is explicitly represented and can be

queried. All information within the ILE platform is represented in the ontology in order to facilitate entity linking and analysis.

The ILE ontology is too large to reproduce it in full in this paper; it comprises 19 high-level domain concepts which are further refined into a total of ~140 concepts and a taxonomy of ~400 specialized relationship types. It has been documented in [8]. The ontology conceptualizes the domain on three levels: meta-level where concept types are captured, the type level, where domain concepts are represented in terms of types, and the instance level, where instance-level data is represented and linked. For example, the meta-level defines EntityType, RelationshipType, and MetaAttributeType. Their instances on the level below represent persons, organizations, (and more broadly a hierarchy of object types), concrete domain relationships that may be established between objects (for example that a Person works for an Organization), and meta-data attributes related to access control, provenance, and temporal validity.

These domain concepts are closely aligned with the draft National Police Information Model (NPIM), complemented with relevant aspects drawn from the NIEM standard[2] and concepts related to case management. The provenance model is an extension of PROV-O [9]. The instances of the domain concepts form the objects comprising the Knowledge Graph on the lowest layer in the ontology. The aforementioned concepts are complemented with classes and objects representing data sources linked to the domain information stored therein as well as schema mapping information required to translate between the external source and the ontology model adopted within the federated architecture.

This multi-level modelling method has been adopted to provide a modular and extensible knowledge representation architecture. The semantic technologies that underpin our platform facilitate incremental addition of elements to the ontology, and phasing out of obsolete concepts can be implemented via meta-data annotations interpreted by the underlying information systems. Changes in information representation received from external parties can be addressed by ontology matching techniques and machine learning methods for information extraction and linking. Profound changes in the information acquisition pipelines however would require changes to the underlying information system. Our modular architecture has been designed to accommodate such changes.

Information from external sources is sought based on a *catalog* of data sources that are available to the system, each with a corresponding *adapter* that communicates with the external systems and rewrites the information and meta-data into the ontology used within the ILE platform [11]. Our platform spans several sources, including an entity database (Person, Objects, Location, Event, and Relations), a case management system, and a repository of unstructured documents.

Information received from external systems is passed through an *ingestion and enrichment* pipeline where entities are extracted [14], enriched with meta-data (provenance and access restrictions) and linked to the knowledge graph in the linked data store.

[2] https://www.niem.gov/

Analytic services include entity extraction from unstructured text [14], entity linking, similarity calculation and ranking. Services provided by commercial tools, such as network analysis and entity liking/resolution solutions, can be integrated in the modular architecture.

Automation services will provide workflow orchestration and alert notices if new information relevant to a case becomes available. Workflow services will facilitate the enactment of work processes such as acquiring authorization and warrants. The automation services component is pending implementation.

Cross-cutting technical concerns, including access control and user management, logging, monitoring and other deployment facilities, have been omitted in this architecture view. Our implementation builds on open source big data technologies (Hadoop/Spark, polyglot persistence, message queues, and RESTful interfaces). The technical building blocks are outlined in [8].

Fig. 1 draws the overall architecture and plot the direction of legal workflow processing.

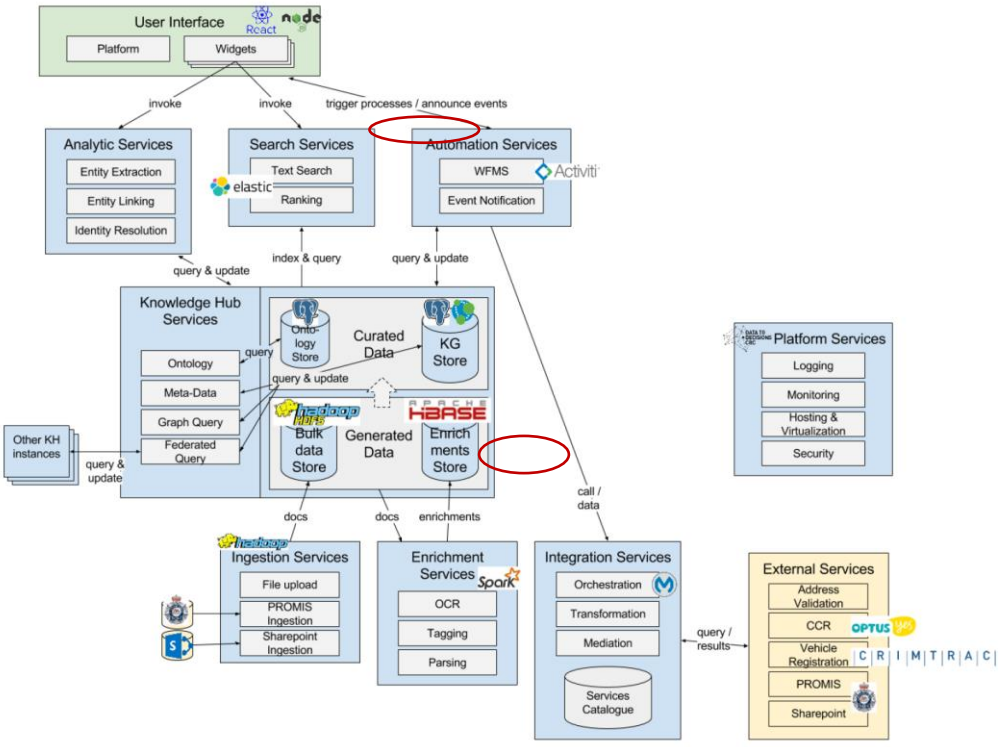

Fig. 1 Architecture overview and legal workflow processing

4 Legal and governance issues

A key concern is the incorporation of legal risk management and compliance constraints into the workflow execution to ensure observance of and compliance with the applicable legal rules, for example agency and privacy rules as well as internal agency policies and procedures. Natural Language parsing can be used to elicit event specifications that could then be translated to business rules in an executable formal language and issued to an event processor in the knowledge hub [16]. These rules would be used to check and guarantee conformance of analytic processes/workflows and data usage. [16] provides support for extracting data from a variety of sources (relational databases, CSV files, JSON, and XML), for modeling it according to a vocabulary of the user's choice and for integrating multiple data sources. This process deserves a closer attention, because (i) it implies LEAs cooperation, and (ii) must be compliant with Australian law.[3]

The D2D CRC's law and policy team outlined for discussion a set of high level principles that may guide the development of an appropriate framework: (i) engender public confidence in government use of data and analytic tools, (ii) develop principles for data governance in National Security Law Enforcement (NSLE) agencies; (iii) employ clear and consistent principles in developing legal frameworks, (iv) improve processes to enhance effective use of data within NSLE agencies, (v) ensure the continued effectiveness of the oversight regime as technologies and NSLE agency practices evolve, (v) disentangle elements of technological change associated with 'Big Data', (vi) maintain data integrity and security in a high volume environment, (vii) ensure fair and appropriate use of data analytics, (viii) use appropriate systems for data matching, data integration or federated access that takes account of benefits and risks; (ix) ensure efficient, appropriate, and regulated sharing of specific data for NSLE purposes [17].

In a recent survey we carried out on the state of the art of Compliance by Design (CbD) [12], we found that the passage from Business to Legal CbD mainly follows a semantic path, in which Natural Language Processing (NLP), non-monotonic defeasible logic and inferential reasoning are combined with enriched annotated legal sources (e.g. described according linked data standards). This is aligned with recent developments in e-business[4] [26] and e-government [27]. Architectures are deemed

[3] We consider primarily investigations conducted by Australian law enforcement agencies, where compliance with Australian laws governing these investigations and subsequent legal proceedings is paramount.

[4] ISO/IEC 42010:2007 defines "architecture" as: "The fundamental organization of a system, embodied in its components, their relationships to each other and the environment, and the principles governing its design and evolution". It has been fleshed out by [26], and [27] [28] for the e-government architecture. See esp. ADM Architecture Requirements Management, and the Architecture Compliance steps defined at TOGAF 9.1, Part VII: Architecture Capability Framework Architecture Compliance. http://pubs.opengroup.org/architecture/togaf9-doc/arch/

to be understandable, robust, complete, consistent and stable. [28] has proposed a comprehensive approach to develop interoperable European e-government services adapting and extending the existing enterprise architecture requirements. The investigation shows that at least half —i.e. not all— of the 30 requirements identified are adequately addressed by enterprise architectures (EA).[5] It concludes with ten interoperability challenges that should be taken into account and addressed when providing pan-European e-government services (PEGS) across Member State borders. Quoting at length: (i) critical success factors should be identified, (ii) an EA framework for PEGS should be built upon widely accepted principles and strategies, (iii) it should comprise architecture design principles and guidelines to reason about alternative design strategies, (iv) in order to facilitate stakeholder management, it should refer to abstract stakeholder classes and roles in interoperability projects and determine drivers for their engagement, (v) the creation of contents can be improved through a methodology that supports the capturing of requirements from business-driven needs, policy implementation processes and other strategic aspects in order to establish common path and to increase the acceptance of architecture outputs among stakeholders (vi) another methodology should describe how to define interoperability specifications on semantic and organizational level, which can be used as a basis for collaboration agreements, (vii) a detailed design of each architecture should identify relevant model fragments and should be based on a commonly agreed architecture description language, (viii) there are missing guidelines and methods that describe how to transition and to govern architectures in multi-stakeholder environments, (ix) several independent implementations of PEGS have to be coordinated, extended and sustained over time (e.g. it should integrate appropriate assessment methodologies to measure specifications and the compliance of solutions with the underlying collaboration agreements, (x) other assessment methodologies can help to determine the level of business standardizations in a domain and to appraise the maturity of market solutions in order to detect appropriate ways forward.

This is a valuable programme. Likewise, we have also devised one close to it with the Australian framework in mind. However, as [28] underline as well, business languages do not completely match all governance and security requirements. Interoperability frameworks do not enable an anticipatory management [29].

Legal compliance is complex, even in relation to national laws where the jurisdiction concerned is a unified, non-federal national state. There are several methodologies and languages to represent norms using formal rules —e.g. Regorous and LegalRuleML [31]—, but there are not *fully* automated ways to carry out such a task. Legal norms must be interpreted in particular fields according to the specific domains to which they apply, anticipating the possible risks and unintended side effects. In addition, ethical principles can nuance or mould this interpretation according to different jurisdictions — e.g. Fair Information Practices in USA, or Data Protection Principles

[5] The 30 requirements obtained in the survey of the literature have been structured into six categories [28]: project management (PM), stakeholder management (ST), service development (SD), interoperability layers and architecture viewpoints (LV), building blocks (BB), and collaboration agreements (CA).

similar to the brand new General Data Protection Regulation in Europe. Similarly, information governance rules and policies differ between private corporations and state agencies.

At a more general level, legal scholars have noticed that the protection of relative civil rights such as privacy does not necessarily entail tradeoffs [21][6]. Nevertheless, as we have already suggested, there are many ways to comply with rule of law requirements, depending on the plurality of legal constraints and constitutional specifications. Apparently, protections for civil rights are not as clear —and arguably as strong— in Australia as in the EU, where the police and their criminal intelligence functions operate subject to well-developed data protection and privacy norms. In contrast to a more comprehensive, integrated EU approach, it could be argued that public transparency and operational secrecy are, for example, not as finely balanced under current Australian law [19]. Contrary to European provisions, the *2017 Australian Productivity Commission Inquiry Report on Data Availability and Use*, excludes national security data [20]. As asserted by the Report, governments use data to monitor and investigate compliance and implement enforcement actions. They retrieve, extract and analyse information from publicly available sources (Open Source Intelligence, OSINT) in a way that can also be regulated [22].

Having a closer look, problems about fragmentation and interoperability are analogous in both Australia and Europe. Different as they might be, the post-facto investigations about the Abdelsam brothers in the Bataclan crisis in Paris [30] and the inquest into the deaths arising from the Lindt Café siege in Sydney[7] have come to similar conclusions. Cooperation among state departments and agencies; and between Law Enforcement Agencies (LEA), can and should be improved.

These conclusions are not limited solely to security issues but can be also extended to the coordination of public administration and the legal system in policy domains. For instance, in many situations, problems might arise "because of gaps of information flow between the family law system, the family violence system, and the child protection system: in many circumstances, important information is not being shared among courts and agencies and this is having a negative impact on victims, impeding the 'seamlessness' of the legal and service responses to the family violence".[8]

Disparity is produced as well across all Australian jurisdictions. At a federal level the *Privacy Act 1988,* for example, regulates the handling of personal information by the federal government and the private sector. The Act does not extend to state governments. Some states have their own comprehensive frameworks. In Victoria, for example, the Privacy and Data Protection Act 2014 (Vic) contains the following Information Privacy Principles (IPPs) which apply to all information held by the Victorian public sector (including the police and a contracted service provider): (i) Open and transparent management of personal information, (ii) Sensitive information, (iii)

---

[6] International human rights law distinguishes between absolute and relative rights. Absolute rights —such as freedom from slavery, torture and servitude— cannot be suspended, restricted, or limited for any reason. Non-absolute or relative rights are those which stand in the various private and legal relations, and can be discussed, re-defined or qualified.

[7] http://www.lindtinquest.justice.nsw.gov.au/

[8] Australian Law Reform Commission (2010), as quoted in [20].

Right to anonymity/pseudonymity, (iv) Notification of collection, (v) Purpose test for use / disclosure, (vi) Direct marketing restrictions, (vii) Cross border disclosure, (viii) Government-related or unique identifiers, (ix) Data quality, (x) Data security, (xi) Access and correction. South Australia, on the other hand, only has an administrative instruction requiring government agencies to comply with a set of Information Privacy Principles while Western Australia does not currently have a comprehensive legislative privacy regime.

Australia has a comprehensive oversight regime in relation to national security and law enforcement agencies. Different bodies have however oversight over different agencies or have oversight over closely-defined aspects of a range of agencies. The fragmented nature of the oversight framework in Australia "will be challenged by an environment where NSLE agencies collaborate more closely in a Big Data framework" [19]. But to reach this milestone, it is our contention that the information integration process that takes place on the platform through reusable ontologies and vocabularies requires a broader regulatory framework. To overcome the patchwork of disparate and sometimes contradictory legal constraints, we will work within an intermediate implementation level, setting what can be called an "anchoring institution" between the semantic tools of the platform and LEAs (end-users).

This set of intermediate conceptual rules constitute a semantic web regulatory model (SWRM), i.e. a specific cluster of guidelines to regulate the information flow, establishing a system of check and balances between LEA's investigative powers and their use of semantic technology [22]. This is an indirect strategy for Compliance by Design (CbD) purposes, in which police officers might set forth internal and external controls, and adopt a conceptual scheme to implement privacy and security principles including ethics as a main component, i.e. at the intermediate level of linked democracy [23]. To encompass both behavioural and informational trends, we use the expression 'Compliance through Design' [CtD] [12] . This means that the increasing of pressure on human compliance management resources in the security area can be taken into account [25]. The crucial point is the coexistence of both artificial and human decision-making and information processes.

Likewise, 'linked democracy' can be defined as "a meso-level approach to both online and offline innovations that elucidates the interactions between people, technology, and data in particular settings, providing a framework of analysis to understand the emerging properties (and tensions) of these interactions" [24]. Therefore, public principles such as transparency, accountability and security could be graduated and connected within particular investigations according to their weight at their specific implementation level. This entails the emergence of different notions, degrees and values of legal compliance, enhancing their semantic side, and outstripping the traditional obstacles of operating from separate information silos.

It is worth noticing that from this pragmatic approach, interoperability does not only mean 'semantic interoperability' —the creation of a common meaning for information ex- change across computational systems— but *systemic* interoperability. That is, the ability of complex systems to interact, share, and exchange information. It focuses onto the coordination of practices, including human behavior, organizational structures, tools, languages, and techniques [23]. Establishing such a model, translat-

ing legal and systemic conditions to institutional and computational constraints and requirements, is the next step.

Acknowledgements

This research was partially funded by the *Data to Decisions Cooperative Research Centre (D2D CRC)*, with participation of the Spanish Project *DER2016-78108-P*. Views expressed herein are however not necessarily representative of the views held by the funders.

References

1. Edwards, M., Rashid, A., Rayson, P. A systematic survey of online data mining technology intended for law enforcement. ACM Computing Surveys (CSUR), 48 (1) (2015): 1-56.
2. Scheepers, R., Whelan, C., Nielsen, I., Burcher, M., Integrated Law Enforcement Project, Qualitative End User Evaluation, Baseline Report. Technical Report, Data 2 Decisions CRC, 2017.
3. Mayer, W., Stumptner, M., Grossmann, G., Jordan A. Semantic Interoperability in the Oil and Gas Industry: A Challenging Testbed for Semantic Technologies. In AAAI Fall Symposium on Semantics for Big Data, Arlington, Virginia, November 2013.
4. Law and Policy Program. Big Data Technology and National Security - Comparative International Perspectives on Strategy, Policy and Law: Australia (Data to Decisions CRC), 2016.
5. Marks, A., Bowling, B. and Keenan, C. Automatic Justice? Technology, Crime and Social Control. In: R. Brownsword, E. Scotford and K.Yeung (eds), The Oxford Handbook of the Law and Regulation of Technology, Oxford University Press, pp. 705-730, 2017. Available at SSRN: https://ssrn.com/abstract=2676154 [2015]
6. Data to Decisions CRC Big Data Reference Architecture, vol. 1-4, Technical Report, Data to Decisions CRC, 2016.
7. Heath, T., Bizer, C. Linked Data: Evolving the Web into a Global Data Space. Synthesis Lectures on the Semantic Web: Theory and Technology, Morgan & Claypool Publishers, 2011.
8. Grossmann, G., Kashefi, A.K., Feng, Z., Li, W., Kwashie, S., Liu, J., Mayer, W., Stumptner, M. Integrated Law Enforcement Platform Federated Data Model, Technical Report, Data 2 Decision CRC, 2017.
9. Lebo, T., Sahoo, S., McGuinness, D., Belhajjame, K., Cheney, J., Corsar, D., Garijo, D., Soiland-Reyes, S., Zednik, S. and Zhao, J. Prov-o: The PROV ontology. W3C recommendation, 2013.
10. Stumptner, M., Mayer, W., Grossmann, G., Jiu, J., Li, W., De Koker, L., Mendelson, D., Bainbridge, B., Watts, D., Casanovas, P. An Architecture for Establishing Legal Semantic Workflows in the Context of Integrated Law Enforcement, Workshop on Legal Knowledge and the Semantic Web (LK&SW-2016), International Conference on Knowledge Engineering and Knowledge Management, Bologna, Italy, Nov. LNAI, 2017 (forthcoming).
11. Bellahsene, Z., Bonifati, A., Rahm, E. Schema Matching and Mapping. Data-Centric Systems and Applications, Berlin, Heidelberg: Springer, 2011. DOI: 10.1007/978-3-642-16518-4.
12. Casanovas, P; González-Conejero, J. Technical Bases for Compliance by Design (CbD). CRC D2D Deliverable, May 2017 (updated July 2017).

13. Suchanek, F.M., Kasneci, G. and Weikum, G. YAGO: a core of semantic knowledge unifying Wordnet and Wikipedia. In Proceedings of the 16th international conference on World Wide Web, pp. 697-706, ACM, May 2007.
14. Del Corro, L., Gemulla, R. Clausie: clause-based open information extraction. In Proceedings of the 22nd international conference on World Wide Web, pp. 355-366, ACM, May 2013.
15. Cardellino, C., et al. Licentia: A Tool for Supporting Users in Data Licensing on the Web of Data. Proceedings of the 2014 International Conference on Posters & Demonstrations Track-Volume 1272. CEUR-WS. org, 2014.
16. Gupta, S., Szekely, P., Knoblock, C., Aman, G., Taheriyan, M., Muslea, M. Karma: A system for mapping structured sources into the Semantic Web. In E. Simperl, P. Cimiano, A. Polleres, Ó. Corcho, and V. Presutti, editors, The Semantic Web: Research and Applications - 9th Extended Semantic Web Conference, ESWC 2012, Heraklion, Crete, Greece, May 27-31, 2012. Proceedings, LNCS 7295, pp. 430-434, Springer, 2012.
17. Law and Policy Program. Big Data Technology and National Security - Comparative International Perspectives on Strategy, Policy and Law: Australia (Data to Decisions CRC, 2016).
18. Casanovas, P., De Koker, L., Mendelson, and David Watts. "Regulation of Big Data: Perspectives on strategy, policy, law and privacy." Health and Technology (2017): 1-15.
19. Bennet-Moses, L., de Koker, L. Open Secrets: Balancing Operational Secrecy and Transparency in the Collection and Use of Data for National Security and Law Enforcement Agencies, CRC Report, 2017.
20. Australian Government. Data availability and use. Productivity Commission Inquiry Report, No. 82, 31 March 2017.
21. Pagallo, U. Online Security and the Protection of Civil Rights: A Legal Overview. Philosophy & Technology 26 (2013): 381–395.
22. Casanovas, P. Cyber Warfare and Organised Crime. A Regulatory Model and Meta-Model for Open Source Intelligence (OSINT). In R. Taddeo and L. Gkorioso, Ethics and Policies for Cyber Operations, pp. 139-167: Dordrecht: Springer International Publishing, 2017.
23. Casanovas, P., Mendelson, D. & Poblet, M. A Linked Democracy approach to regulate health data. Health and Technology (2017), DOI: 10.1007/s12553-017-0191-5
24. Poblet, M. and Plaza, E. Democracy Models and Civic Technologies: Tensions, Trilemmas, and Trade-offs, 2017, *arXiv preprint arXiv:1705.09015*.
25. Watts, D., Bridget Bainbridge, B., de Koker, L., Casanovas, P., Smythe, S. Project B.3 A Governance Framework for the National Criminal Intelligence System (NCIS), Data to Decisions Cooperative Research Centre, La Trobe University, 30 June 2017.
26. Open Group Standard TOGAF Version 9.1 Document Number: G116. ISBN: 9789087536794.
27. Brous, P., Janssen, M., Vilminko-Heikkinen, R. Coordinating Decision-Making in Data Management Activities: A Systematic Review of Data Governance Principles. In H. J. Scholl et al. International Conference on Electronic Government and the Information Systems Perspective EGOVIS 2016, LNCS 9820, pp. 115-125, Springer International Publishing, 2016.
28. Mondorf, A., Wimmer, M.A. Requirements for an Architecture Framework for Pan-European E-Government Services. In H. J. Scholl et al. International Conference on Electronic Government and the Information Systems Perspective EGOVIS 2016, LNCS 9820, pp. 135-150, Springer International Publishing, 2016.

29. Mondorf, A. and Wimmer, M., Contextual Components of an Enterprise Architecture Framework for Pan-European eGovernment Services. In *Proceedings of the 50th Hawaii International Conference on System Sciences*, HICSS, 2017, pp. 2933-2942.
30. Bureš, O., 2016. Intelligence sharing and the fight against terrorism in the EU: lessons learned from Europol. European View, 15(1): 57-66.
31. Sadiq, S., Governatori. G. A methodological framework for aligning business processes and regulatory compliance. In: Jan van Brocke and Michael Rosemann, editors, Handbook of Business Process Management, pp. 159-176, Springer, 2010.