Artificial Intelligence and Law

A Trust and Governance Architecture for Information Sharing across Domains

--Manuscript Draft--

Manuscript Number:	
Full Title:	A Trust and Governance Architecture for Information Sharing across Domains
Article Type:	Original Research
Keywords:	Information sharing; trust; governance
Corresponding Author:	William J Buchanan, PhD Edinburgh Napier University Edinburgh, UNITED KINGDOM
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Edinburgh Napier University
Corresponding Author's Secondary Institution:	
First Author:	William J Buchanan, PhD
First Author Secondary Information:	
Order of Authors:	William J Buchanan, PhD
	Omair Uthmani, PhD
	Lu Fan, PhD
	Burkhard Schafer, PhD
Order of Authors Secondary Information:	
Abstract:	This paper outlines a novel architect which integrates three main components: a trust framework, governance rules, and a gateway, in order to implement information sharing across domains. A key element of this is the modelling of the trust policies that exist between domains, thus the paper outlines a novel method, using Binary Decision Diagrams (BDDs), to model these.

Prof Bill Buchanan

Bill Buchanan is a Professor in the School of Computing at Edinburgh Napier University, and a Fellow of the BCS and the IET. He currently leads the Centre for Distributed Computing, Networks, and Security, and works in the areas of security, next generation user interfaces, Web-based infrastructures, e-Crime, intrusion detection systems, digital forensics, e-Health, mobile computing, agent-based systems, and simulation. Bill has one of the most extensive academic sites in the World, and is involved in many areas of novel research and teaching in computing. He has published over 27 academic books, and over 130 academic research papers, along with several awards for excellence in knowledge transfer, and for teaching, such as winning at the I ♥ my Tutor Awards (Student voted), Edinburgh Napier University, 2011, and has supervised many award winning student projects.

Presently he is working with a range of industrial/domain partners, including with the Scottish Police, health care professionals and the finance sector. As part of the drive to create a World-leading infrastructure for security and cybercrime, he leads the Scottish Centre of Excellence for Security and Cybercrime which brings together a wide range of collaborators, including most of the universities in Scotland, the Scottish Police, the public sector, and a range of SMEs and large organisations. Current work includes initiatives on creating an e-Forensics Cloud across Scotland, and in organising a large-scale Symposium to engage a large number of stakeholders within Scotland, in order to focus on creating a World-leading infrastructure.

He has a long track record in commercialisation activities, including being a co-founder of Inquisitive System, which has progressed from PhD work to a university spin-out, though the Scottish Enterprise funded Proof-of-Concept scheme. This spin-out has also involved patenting novel security software in three territories around the World. His current work includes collaboration of TSB Grants with Microsoft plc on a £2million project which aims to improve the care of the elderly using Trusted Cloud-based services, and with Chelsea and Westminster Hospital on a next generation Health Care platform. This also matches up with other funded projects with the FSA and the Scottish Police.

Dr Lu Fan

Dr. Lu Fan received the BSc (2003) in Computer Networking from Nanjing University of Science and Technology, Nanjing, China, and the MSc (2005) with distinction in IT Software Systems and the PhD (2009) in Computer Science from Heriot-Watt University, Edinburgh, UK. Currently he is a Senior Research Fellow of Centre for Distributed Computing, Networking and Security, School of Computing, Edinburgh Napier University, Edinburgh, UK. He is the Technical Lead of multiple award winning research projects, including DACAR, Trusted Service, Cloud4Health and sa.FIRE. He designed and implemented a Cloud-based e-Health platform, which reinforces the integrity, security, confidentiality and auditability of medical data, and facilitates the development, integration, and cost-effective delivery of e-Health services. Furthermore, Dr Lu Fan is a member of the Technical Program Committee of International Conference on eHealth, Telemedicine, and Social Medicine (eTELEMED).

Prof Burkhard Schafer

Burkhard studied Theory of Science, Logic, Theoretical Linguistics, Philosophy and Law at the Universities of Mainz, Munich, Florence and Lancaster. My main field of interest is the interaction between law, science and computer technology, especially computer linguistics. How can law, understood as a system, communicate with systems external to it, be it the law of other countries (comparative law and its methodology) or science (evidence, proof and trial process).

As a co-founder and co-director of the Joseph Bell Centre for Legal Reasoning and Forensic Statistics, he helped to develop new approaches to assist lawyers in evaluating scientific evidence and develop computer models which embody these techniques. A special interest here is the development of computer systems that help law enforcement agencies to co-operate more efficiently across jurisdictions, assisting them in the interpretation of the legal environment within which evidence in other jurisdictions is collected. This research is linked to my wider interest in comparative law and its methodology, the idea of a "Chomsky turn in comparative law", and the project of a "computational legal theory"

He is involved with a number of organisations that promote the exchange between computer science and law, including the German Association for Informatics, BILETA, and the Evidence and Investigation network of the Scottish Institute for Policing Research. He’s also on the Nomination Committee of the International Association for Artificial Intelligence and Law.

Burkart is currently the Director of the SCRIPT Centre for IT and IP law, working mainly on issues such as privacy compliant software architecture and more generally the scope and limits of representing legal concepts directly in the internet infrastructure.

Interagency Data exchange, Privacy protection and Governance Architecture for Information Sharing across Domains

Abstract. This paper discusses privacy enhancing technologies in the field of inter-agency data sharing, a key government objective for efficient public service delivery. We analyse the legal and conceptual framework that governs multi-agency cooperation in particular in the field of child protection work, and develop two approaches to represent relevant data protection ideas computationally in the infrastructure that agencies use to exchange sensitive personal data.

Keywords: Information sharing, trust, governance

1 Introduction

1.1 Data sharing, data privacy and the UK e-governance agenda

This paper describes a new approach to privacy protection by design in the field of inter-agency data exchange. Its aim is to assist the legitimate interest of state agencies to share data in cases such as multi-agency child protection work, while at the same time assuring both individual citizens and the agencies that hold their data that any sharing of it complies not just with the letter, but the spirit of data protection law. It grew out of interdisciplinary research facilitated by the Scottish Institute for Policing Research (SIPR) that brought together, lawyers, social- and computer scientists to work with data protection officers in police, social and health care services. We first introduce the guiding philosophy of our approach buy locating it in the wider socio-legal discussion on privacy and public service delivery in the UK, looking in the influential conceptual analysis by Perri, Bellamy and Raab (henceforth PBR). From their analysis, we take two ideas in particular: that in designing IT infrastructure for privacy compliant data sharing in the public sector, it is helpful to distinguish horizontal vs vertical integration. Horizontal data sharing policies, in the terminology of PBR, are those policies, laws and guidelines that regulate the tension between data sharing and privacy protection globally, across all services and applications. They are the most general and therefore abstract provisions that we encounter to determine if a specific data sharing event was law compliant. Vertical policies by contrast involve regulation and rules for very specific purposes, agencies and data sharing agreements. They tend to be as a consequence highly specific, technical and detailed. For Privacy Enhancing Technologies which in the spirit of Larry Lessig try to encapsulate legal provisions directly in code and making the law thus “self-enforcing”, horizontal and vertical policies pose very different challenges. Horizontal policies will be very abstract, using as all high level legal provisions vague terms, policy statements and permit considerable discretion in their application. For a formal rendition, this raises the well-known and frequently discussed issue in AI and law of how to model open textured rules. (see e.g. Bench Capon 1993; Stranieri et al 1999) . Vertical rules will be highly specific, and therefore in constant need to be updated and revised as circumstances change. This raises the equally often discussed problem of consistency maintenance in legal AI in the face of rapidly changing laws (cf e.g. Bradley et al 1991; Boer, Engers and Winkels 2010). PET approaches will have to take the different type of legislation in horizontal and vertical settings into account, and idea represented in the structure of this paper. In the first part of the technical discussion, we show how we can model the balancing between privacy and data sharing requirements in a specific domain – multi-agency cooperation for law enforcement purposes. In the next part, we abstract from this a novel approach to a privacy enhancing data sharing framework globally, across the public sector, a horizontal approach in PBR's terms. Another insight we take form PBR's study is the important of "trust" if a non-arbitrary "balancing " between privacy and data sharing is to be achieved. This in our analysis applies not just to the relation between citizen and state, but also as trust relation between different state agencies, in particular agencies with widely differing professional cultures such as social services and police. Before we discuss the technical solutions however, we will situate the discussion more broadly in the socio-legal debate on surveillance, data sharing and privacy protection.

1.2 Situating the data sharing, and data privacy debate in the UK

In their seminal study on Data protection and information sharing in Britain, Perri, Bellamy and Raab (2004a and 2004b) analyse the difficulties in managing the tension between the increasing commitment of the UK government to interagency data sharing, and the emerging, more privacy friendly legal framework driven by the Human Rights Act and the European Data Protection Directive, implemented through the Data Protection Act. While their analysis focussed on policies by the Labour government that was replaced by a conservative administration in 2010, the same policy trajectory has largely been maintained, even though the enthusiasm for very large and centralised IT projects in the delivery of data sharing has become more muted, with several high profile projects such as the national ID database for the time being abandoned (Schafer 2011).

Perri, Bellamy and Raab identify four possible strategies for the governance to deal with the relation between data sharing and privacy. Governments, through a combination of legislation and IT infrastructure, could either

a) seek to make the commitment to efficient public service delivery through data sharing and the commitment to privacy protection consistent;
b) mitigate the tensions with safeguards such as detailed guidelines;
c) allow privacy to take precedence over integration;
d) allow data sharing to take precedence over privacy.

The potential conflict between data sharing and privacy protection that these four strategies address exist in principle independently of developments in computer technology. However, as several commentators have noted, advances in ICT and the sometimes aggressive promotion of e-government by subsequent governments that have committed themselves to a huge extension of online service provision have increased the tension (Bellamy 1999, Beynon-Davis 2007). Data sharing is also a prerequisite to deliver key social policy programmes, in particular "holistic, multiagency" interventions that focus on small neighbourhoods or target specific groups or even individuals (Perri Bellamy and Raab 2004a), perceived to be either "at risk" (e.g. policies to combat school truancy) or "a risk" (fight against social ills ranging from anti-social behaviour by juveniles to terrorism). The risk management paradigm with its ever elusive promise of "safety" (Rauhofer 2008) and an actuarial approach to policy areas ranging from medical services to crime prevention further contributes to the enhanced reliance on big data for government decision making. These approaches require a high degree not just of data, but data of high quality and integrity, to give legitimacy to increasingly discriminating decision making in areas such as welfare entitlement, medical support or formal sanctions. To implement this agenda, a report from the Cabinet Office's Performance and Innovation Unit (the PUI, as it then was) recommended a new legal framework that increased considerably the powers to share information about individual citizens across public service agencies. This was softened by suggestions to develop new and better techniques and computing infrastructure for better privacy protection (PIU 2002). Even though the PIU report was not implemented in its original form, it illustrates, as PBR show, the possible tensions between data sharing and privacy in the field of public sector service delivery. The policy imperatives that make data sharing highly desirable, - risk assessment, data matching and social sorting techniques – are intrinsically linked with an increased capacity by the state for surveillance (Gandy 1993; 6 1998; Lyon 2003). At the same time, the high profile incorporation of the European Convention of Human Rights, whose Article 8 protects a right to privacy, and the increasing importance of European Union approach to Data Protection provided an environment where the government also committed itself publicly as a champion of individual liberty against state intrusion.

PBR conclude that the government's main strategy to resolve this tension has been so far b): more and more, and more and more detailed, rules, guidelines and regulation for highly specialised applications that try to mitigate the default position that data between government agencies ought to be shared. This makes data protection often cumbersome in practice, with businesses and state agencies complaining about a high administrative burden with little or no actual benefits for data subjects, and detrimental consequences for service delivery. (see e.g. Choudrie, , Vishanth and Jones 2005) PBR's assessment is consequently sceptical: This strategy is potentially unstable and possibly unsustainable. In particular, there is a very pressing concern that data sharing is taking precedence over privacy, with the specific legal safeguards inefficient or becoming irrelevant. While PBR thus give voice to the concerns of many academic analysis and privacy advocates, as noted above, the opposite picture is more frequently found in the public discourse in the UK. There, privacy laws are often portrayed as either a way "for criminals to get away with it", hampering legitimate police efforts, or as an intrinsically undemocratic measure that protects councils and government agencies from public scrutiny by limiting transparency.

Despite these problems in the practice of e-governance, PBR also note that data sharing and privacy need not necessarily be in conflict. This can be seen when comparing different government strategies in different sectors, with the balance between privacy and data sharing requirements differing markedly between applications and agencies, as does the perceived degree of conflict between the two. To support this claim, PBRs's analysis stipulate two distinct, yet interdependent levels on which the resulting tension is managed. In what they call the horizontal dimension, a negotiated compromise is attempted at the level of general data protection law and the rules that govern data sharing practices across the public sector. Secondly, they identify as "vertical dimension" the attempts to "balance" the conflicting demands that are specific to particular fields of public policy and service delivery. As examples of vertical approaches to balance data sharing against privacy concerns, they analyse three different policy fields - Crime reduction and public protection; Data sharing and data matching to reduce welfare fraud; and Data sharing in the National Health Service. These case studies support their conclusions that:

> "public bodies are struggling to reconcile imperatives for data sharing with the principles of data privacy and that there is, as a result, an incipient tension between joined-up government and the right to privacy" (PBR 2004 b p.411)

They also notice however that the case of the NHS differs markedly from the other examples, with the management of the tension becoming a focal point of its information-governance regime. This has resulted in particular tight and demanding rules put in place to minimise the risk from data sharing. In the case of crime and public protection by contrast, and even more in the chase of fighting benefit fraud, the government allows data protection principle principles such as tight restrictions on use or avoidance of excessive collection, to take back seat behind the overriding imperative for joined-up working.

In what follows, we will develop ideas for a privacy- preserving, information-sharing infrastructure that takes PBRs analysis as their conceptual starting point, fleshed out however with empirical studies, both from the academic literature, our own work with key stakeholders in the quest for efficient inert-agency data sharing, and a questionnaire based study we conducted for a subgroup amongst these shareholders. The overall aim of our approach is to show how strategy1, co-existence between data protection and efficiency demands, desirable in theory but so far difficult to achieve in practice, can be facilitated through Privacy Enhancing Technology that embeds privacy concepts directly into the information sharing infrastructure, ensuring as a result data protection by design. The "glue" between the two concepts is the notion of trust: As PBR notice, the government has argued that good data protection makes a positive contribution to the level of trust between agencies, and also between agencies and their clients or customers. (PBR 2004a p216). This aligns with our own experience working with multi-agency stakeholders. Inherited, traditional distrust between different government agencies can be a serious barrier to information sharing, having sound and robust data protection mechanisms in place can mitigate the hesitance to share data. This applies to situations where citizens are asked to supply data because they personally will benefit from better services (e.g. allowing different hospitals to access my health data so that I can chose the one best suited for my needs), but also to situations where the gathering of data does not benefit any individual citizen, but serves communal interests, such as data sharing between police and education sector to minimise truancy.

We will begin with a case study of data sharing in a specific domain, crime prevention and control, focussing in particular of data sharing in multidisciplinary teams, involving e.g. police, social services, medical services and education providers. This cuts across two of PBRs case studies, though they briefly mention child protection work under the crime and disorder header. However, it should be noted that the involvement of medical care providers at a central point in detecting child abuse means that this issue can't be separated from the treatment of data in the NHS environment, a n environment that PBR did not only analyse separately, but which according to them has significant differences in the way in which it mediates between data sharing and privacy objectives, finding a framework that allows privacy compliant data sharing between police and medical professionals is therefore a particularly difficult task. is not only separate, but also Starting with a "vertical", domain specific example of managing the tension between privacy and data sharing follows PBRs conclusion that

> "insofar as formal rules and norms are emerging to manage the tensions, the most important rules and norms are specified vertically, in the policy fields. By contrast, the horizontal policy initiatives, both in data protection law and in cross-cutting data-sharing policy, have limited capacity either to constrain or to direct policy and practice."

In the final part, we will however generalise these ideas to a platform for data sharing independently of domains. The framework we use for this has been tested in a hospital environment – if it can serve as a blueprint, then following the discussion above, the sector with the most stringent data protection requirements would determine the architecture for data sharing in general.

1.3 Vertical data sharing policies: Data sharing and privacy protection in multi-agency crime prevention

In 2005, the 17 month old Peter Connolly, known in the press as "Baby P" died from more than 50 injuries that he suffered over an eight-month period on the hands of his mother and her boyfriend. It quickly transpired that he had been seen frequently by Haringey Children's services and NHS health professionals, who had failed however to coordinate their various reports and as a result spot the danger he was in. His was but the last I a number of high profile cases of child neglect and child abuse where victim and perpetrator had been known to several agencies, from social services to hospitals to police, but where due to a lack of data sharing between them, appropriate reaction had not been taken.

At the same time however, the opposite problem also grabbed headlines: Local councils were caught abusing legislation intended to combat terrorism to collect and exchange data of citizens suspected of everything from permitting their dog to foul in parks to lying about their address in applications to schools for their children.

When in the first type of cases, harm ensued because data that should, and legally could have been exchanged between agencies wasn't, in the second type of cases data that should never have been collected in the first place was exchanged without care and precaution between agencies. In theory, a whole raft of legal measures, from the Data Protection Act to the Community Care Act 2003 should have ensured that all and only the necessary and legally permissible data is exchanged. However, regulating data exchange between agencies through legal codes has proven difficult. "Top Level" acts such as the Data Protection Act use highly abstract language, give only vague guidance on how to balance competing interests, and are too unspecific to be of direct help to personnel that is not legally trained. Inter-agency Information Sharing Agreements that try to operationalise the relevant law have by now become so complex, long and technical in their attempt to cover every possible situation that they are often ignored by practitioners working under sever time constraints. Technological solutions have not fared better. The UK government invested heavily in "one big database" schemes that obligates agencies to store their information centrally. If all the data is in one place, in a uniform format and accessible to everybody who needs to know, so the reasoning, information exchange must improve. Yet the results of these schemes is disappointing so far – not just because they are overly expensive, but because they are faced with a grassroots boycott by users. If the users are forced to record data in ways that are not aligned with their own understanding of their role, and if they furthermore have to fear to lose "ownership" over the data that is stored outside their control, uptake of the technology will be limited. The result is that decisions about data sharing are often done informally, between individuals that trust each other on a personal level, and with insufficient transparency and audit possibilities. Personal attitudes and professional mentalities, rather than legal rules, decide if agencies "play it safe" by not disclosing important information (fearing actions under the DPA), or disclosing unnecessarily (in fear of being caught "doing nothing").

The exchange of information between the police and community partners forms a central aspect of effective community service provision. In the context of policing, a robust and timely communications mechanism is required between police agencies and community partner domains, including: Primary healthcare (such as a Family Physician or a General Practitioner); Secondary healthcare (such as hospitals); Social Services; Education; and Fire and Rescue services.

Such requests typically form the basis for any information-sharing agreement that can exist between the police forces and their community partners. It defines a role-based architecture, with partner domains, with a syntax for the effective and efficient information sharing, using SPoC (Single Point-of-Contact) agents to control information exchange. The application of policy definitions using rules within these SPoCs is inspired by network firewall rules and thus defines information exchange permissions. These rules can be implemented by software filtering agents that act as information gateways between partner domains. Roles are exposed from each domain to give the rights to exchange information as defined within the policy definition. This work involves collaboration with the Scottish Police, as part of the Scottish Institute for Policing Research (SIPR), and aims to improve the safety of individuals by reducing risks to the community using enhanced information-sharing mechanisms. Agencies are actively encouraged by governments (Police and Crime Standards Directorate 2007) to form partnerships and collaborate to ensure provision of effective community services. Working in partnership by sharing information has been particularly successful in public services (Clarence and Painter 1998). Often, partnership working is a requirement mandated by legal directives. In the UK, for example, Acts of Parliament such as the Health and Social Care Act 2001, Police Reform Act 2002, Community Care Act 2003 and the Children Act 2004 all necessitate information sharing among partner agencies.

Barriers to forming effective partnerships and information exchange include lack of trust between organisations; lack of understanding of policies and legislation; and disparate communication systems. The issue of trust can arise from traditional rivalries between organisations that view each other as competitors rather than collaborators (Hudson et al 1999). Our research however indicates a more pertinent problem caused by incongruent professional values and missions of the different stakeholders. A social worker, trying to establish a trust relation with a young person deemed at risk will be hesitant to pass on information about low level drug dealing to the police, if he fears that the information will result in heavy handed police activity that would make his work impossible (Willem and Buelens 2007). However, evidence suggests that increased government encouragement to collaborate (Richardson and Asthana 2006), in the form of incentives and legal obligations, has helped in alleviating this situation. Initiatives that highlight best practices and procedures, such as the guidance on the Management of Police Information (MoPI) within the Scottish and other UK police services, also simplify the interpretation of policies and legal requirements. This ease of interpretation of policies, in turn, alleviates the risks agencies face from non-compliance and, thus, further aids collaboration.

1.4 Data sharing model

In modern democracies, rightly suspicious of the danger that information can also be abused, data sharing has to take place within tightly defined legal parameters, found in legislation such as the Data Protection Act (1998) in the UK. Put simply, it is more acceptable to invade the privacy of a person under reasonable suspicion to plan a terrorist attack than that of a mother suspected, with little evidence, to have lied about her address on the application for a school place. As a first step to model the legally required balancing, we developed four categories of data sharing scenarios that can be found in police work. On each level, different arguments count for or against sharing of data, and the legal analysis differs accordingly:

- **Level 1. Community**. This level focuses on community actions, typically using Intelligence Lead Policing, where measures are taken to try and prevent future criminal activities. A typical example is a decision to increase patrols in an area where intelligence indicates that gang activity could otherwise rise.
- **Level 2. Preventative intervention**. This level focuses on prevention of specific, identified criminal activities, with the requirement to share information often depending on the anticipated harm. A typical example is rescuing a kidnap victim.

- **Level 3**. **Crime investigation**. This level deals with the investigation of a specific crime. Unlike level 2, which is forward looking, level 3 is backwards looking, a singular past event is the focus, the main harm has already occurred.
- **Level 4. After the event**. This level focuses on consuming data on criminal activity, in order that it can be used in the future to reduce the risk to the public. This involves for instance the compilation of statistics by police agencies. It feeds back into level 1, and also informs activities such as resource allocation by the police.

The justification to share information at Level 2 and 3 can be achieved through an information sharing agreement. One key feature of our approach is the ability to formally define the relevant criminal contexts explicitly. For example in a missing person context, a social worker may request the current location of the person from the police, and justify it in this context. When audited, the social worker would then have to provide evidence that the context was correct at the time of the query. A rule can thus be written which defines the context, and the requirements for the information sharing, which is then agreed between the police and the community partner. The information sharing agreement (ISA) at this level can thus define a **criminal context**. At Level 1 and Level 4, it is more difficult to define clearly a criminal context, as there is no actual crime. In this paper we focus therefore on Levels 2 and 3, as the criminal context is easier to define in an ISA.

2 Information Sharing Framework

As discussed in our analyses of PBR, highly specific and local data protection rules and guidelines are typical for vertical data sharing and data protection management. . As a result, the syntax proposed in this part of the paper mirrors principles of best practice within the Scottish Police, such as those highlighted in the guidance on the Management of Police Information (MoPI). They are a typical example of what PBR described as “vertical” rules that manage their tension between privacy and data sharing. These guidance principles for police information management in turn are loosely inspired by concepts such as “firewall” and “single point of contact” that are attuned to computational thinking.

Once the need to share information with a partner agency is identified and affected procedures and compliance issues defined, the principles highlighted in MoPI can be used to construct an Information-Sharing Agreement (ISA). ISA’s define the agreed specific rules, derived from policies, that direct the recording, access, review and dissemination of information between partner agencies. Usually, agencies that have similar functions also have similar ISAs and can be grouped together into domains. From a Scottish policing perspective, common information sharing domains include Police services (POL), Social Services (SOC), Primary healthcare (HCP), Secondary healthcare (HCS), Education (EDU), and Fire and Rescue (FIRE). MoPI also outlines the concept of a Single Point-of-Contact (SPoC), which describes the individuals who are designated as main contacts for the exchange of information between domains.

Any exchange of information between the domains, therefore, needs to occur through the designated SPoCs.

2.1 Single Point of Control (SPoCs)

Figure 1 illustrates the Single Point of Contact (SPoC) concept described in the guidance on the Management of Police Information (MoPI), which is implemented as software agents that serve as gateways for information requests. The function of these SPoC agents is inspired by firewalls within a computer network. At a basic level, firewalls use a defined set of rules to either permit or deny network traffic. Similarly, SPoC agents validate requests for information exchange based on rules, derived from organisational policies and legislative requirements, as defined in Information-Sharing Agreements (ISA). This means that the agent attempts to match a request for information exchange against the rules defined in the set of rules in the ISA. If the request does not match a rule, the agent will then attempt to match the request against the next rule and so on. Once a match is found, the agent will carry out the action (permit or deny), as defined by that rule, and end the searching (as a firewall would). If no matching rule is found in the set, the agent will deny the request. This is similar to the idea of an implicit deny criterion used by firewalls where no matching rule is found. In the case that a request is denied, the agent will return information indicating the reason for the denial. The policies defined in the ISA can take the form of restrictions such as limits on the number of search items returned, specified timeframe of validity for an incoming request, and so on. An agent based approach has the benefit that each partner remains owner of their data and decides its formats, a key requirement to ensure acceptance of the approach and maintenance of trust. Just as in the offline world, a police officer would have to contact his counterpart in another agency, which then determines internally what information to release, the software agents are in constant negotiation with each other for access. This maintains the intuitive concept of ownership of data, and auditable exchange relations. It also permits, through a federated Identity Management System, the protection of the identity of the source of information where this is desirable. We mentioned above the case of a social worker who has information that in principle is relevant for the police, but would force him to put his sources in jeopardy. This approach here permits a nuanced response to an information request, with strong protection of identity build into the system. This too we see as a necessary design feature to increase trust through architecture.

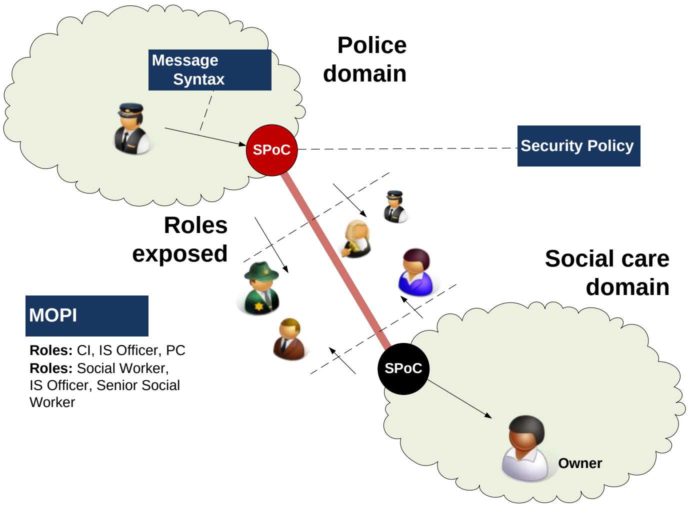

Figure 1: Overview of the architecture

2.2 Role-Based System

A core part of the Information-Sharing Agreement (ISA) is to specify those who will have access to the information. Typically, this involves identifying functional roles that need to access information in order to complete a defined task or job. The information exchange syntax thus uses a hierarchy within domains and roles exposed between domains to facilitate the exchange of information. For example, Analyst (ANA) may be an exposed role from the Child Abuse Investigation (CAI) organisational unit in the Police domain (POL). This role is represented as POL.CAI.ANA, illustrating the full hierarchy. Similarly, an Inspector (INS) from the Missing Persons (MPR) business area of the Police domain would then be represented as POL.MPR.INS. For a Social Worker (SW) role exposed from the Children Day Care Service (CDC) of the Social Services (SOC) domain, the representation would be SOC.CDC.SW. Essentially an exposed role is one that has permissions for information exchange from another domain. For example, if Social Workers (SOC.CDC.SW) are allowed to request information from Police (POL), then the SOC.CDC.SW role would be defined in the ISA as having permissions for this action. Thus, the SOC.CDC.SW role is exposed from the Social Services domain to the Police domain. Crucially again, the definition of these roles happens through each community partner, and is not dictated from the top. This means that traditionally highly hierarchical organizations such as the police can nonetheless efficiently interact with much flatter hierarchies found in social services, without either having to compromise their self-understanding and way of doing things.

2.3 Syntax

A syntactic approach to the concept of information-exchange simplifies the creation and implementation of rules. The main reason for this approach is the vast number of disparate information systems that various police divisions and partner agencies use, and which this approach preserves. The danger is that valuable semantics can be lost in the exchange, which degrades the efficiency of the information-sharing mechanism. By agreeing not to harmonize the way data is stored, but the way information requests are handled, we try to minimize the dangers while retaining the benefits of a decentralized approach. Common logical definitions, which constrain possible interpretations of any given concept to a finite set, therefore, need to be agreed upon before communication can occur

Adding key security elements to this structure yields the proposed syntax for policy rules which are applied into the SPoC:

> [permit | deny] [Requester] requests [Attribute] of [Object] with [Context] from [Owner] for [N] records in [TimeWindow] using [Compliance]

A similar matching syntax can then be applied to the request messages:

> [Requester] requests [Attribute] of [Object] with [Context] from [Owner] within [Start] to [End]

For our purpose relevant elements of this syntax are defined as:

- **[permit | deny]** This part of the rule syntax indicates the action of the rule and defines whether a message meeting the rule criteria will be permitted or denied.

- **Requester** This identifies an exposed role defined in the ISA. For example, this role might be General Practitioner (FAMDOC) in Primary healthcare (HCP) or a Detective Constable (DETCST) in Police services (POL) domain.

- **Object**. This refers to any entity about which information is held, including people, vehicles, events and so on. It is a freeform field.

- **Attribute**. This is a unit of information describing an Object. Attributes may include details about location (address, mobile phone tracking), identity (name, insurance number), history (prior convictions, documented allegations), behaviour (calm, violent) and association (group memberships, known associates).

- **Context**. This identifies the reason why the information is being shared. The context also governs the level of access and permissions associated with information exchange and, hence, affects the priority accorded to information requests. For example, the Emergency context signifies a threat to life or threat of violence and will require a higher priority allocation than a Vandalism context.

- **Owner**. Defines a role with sufficient privileges to manage all aspects of an information element. The owner has the authority to allow or deny access to an information element, as required by legislation and defined responsibilities. Use of the term owner in this context implies custodianship.

- **[Compliance]** This is part of the rule syntax that refers to policies and legislative requirements that affect the exchange of information. Such as the Data Protection Act, the Human Rights Act, the Freedom of Information Act, and so on.

2.4 Context

A key novelty in the proposed system is the use of context for a request, where the ISA will define rights based on the context of the request. For example the rights to data will be higher within the context of a missing persons query than for a trivial access to data. It is thus important that the context levels, and associated rights, are clearly defined in the ISA. For our approach, we developed a conceptual hierarchy loosely based on the categories found in the codified, and hence highly conceptual, German Criminal law. In addition, we use as a proxy to weight severity within a category (e.g. murder vs manslaughter as “offences against the person”) the minimum punishment that the crime carries (Francis, Soothill, and Dittrich 2001), supplemented by a large questionnaire based study that asked members of relevant agencies to rate their own perception of the seriousness of certain offences, and what privacy risk they think is acceptable to either prevent or prosecute such an offence (for details of this study see Uthmani et all 2011) In the next step, socio-legal literature together with our own studies is used to lay an empirical foundation for the metrics the system is using. This additional measure will help us to model one of the main problems in interagency collaboration, diverging value systems that shape professional cultures.

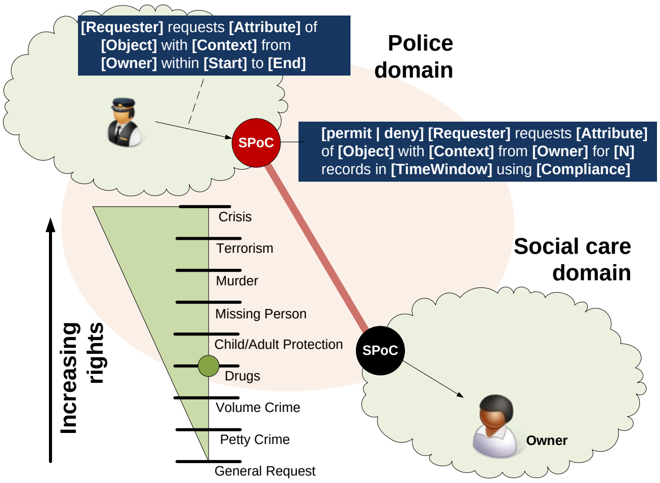
[Requester] requests [Attribute] of [Object] with [Context] from [Owner] within [Start] to [End]
Police domain
SPoC
[permit | deny] [Requester] requests [Attribute] of [Object] with [Context] from [Owner] for [N] records in [TimeWindow] using [Compliance]
Crisis
Terrorism
Murder
Missing Person
Child/Adult Protection
Drugs
Volume Crime
Petty Crime
General Request
Increasing rights
Social care domain
SPoC
Owner

Figure 2: Context definition

2.5 Example

Rules may be used to explicitly permit or deny information exchange requests made by an exposed role. For example, a Senior Family Physician (Requester role=FAMDOCSEN) in Primary healthcare (Requester domain=HCP) is allowed to request a person's medical test results (attribute=MEDTST), from a Laboratory (Owner organisational Sub-unit=LAB) located in a Hospital (Owner organisational unit=HOSP) in Secondary healthcare (Owner domain=HCS), where the person (Object=PERSON) is a patient (Context=PATIENT). A Junior Family Physician (FAMDOCJUN) role from the same domain is not allowed to request this information. These information exchange policies can be used to derive an explicit permit rule (Rule 1) for the FAMDOCSEN role and an explicit deny rule (Rule 2) for the FAMDOCJUN role. These rules would be defined in the Information-Sharing Agreement (ISA) and processed by the SPoC agent (where [PERSON] will be the free-form search field):

Rule 1: [permit] [HCP.FAMDOCSEN] requests [MEDTST] of [PERSON] with [PATIENT] from [HCS.HOSP.LAB] for [N] records in [TimeWindow] using [Compliance]

Rule 2: [deny] [HCP.FAMDOCJUN] requests [MEDTST] of [PERSON] with [PATIENT] from [HCS.HOSP.LAB] for [N] records in [TimeWindow] using [Compliance]

Given the above rules, the following requests may be considered:

Req. 1: [HCP.FAMDOCSEN] requests [MEDTST] of [PERSON] with [PATIENT] from [HCS.HOSP.LAB] within [Start] to [End]

Req. 2: [HCP.FAMDOCJUN] requests [MEDTST] of [PERSON] with [PATIENT] from [HCS.HOSP.LAB] within [Start] to [End]

Thus, a request made by a Senior Family Physician (Request 1) would match Rule 1 and be permitted by the SPoC agent. A similar request made by a Junior Family Physician (Request 2) would match Rule 2 and be denied by the SPoC. In the case of Request 2, the SPoC may return the following message:

Junior Family Physician role does not have permission to access the requested resource.

The context of a request for information exchange affects how the request is handled. For example, a Detective Constable (Requester role=DETCST) in the Domestic Violence (Requester organisational unit=DOM) area in Police services (Requester domain=POL) is allowed to request a person's (Object=PERSON) behaviour information (Attribute= BEHAVIOUR) from the Rehabilitation Support organisation (Owner organisational unit=REHAB) in Social Services (Owner domain=SOC), if this is in relation to a domestic violence investigation (Context=DOM.INVST). This following rule may be derived from this policy:

Rule 3: [permit] [POL.DOM.DETCST] requests [BEHAVIOUR] of [PERSON] with [DOM.INVST] from [SOC.REHAB] for [N] records in [TimeWindow] using [Compliance]

Thus, the following request, Request 3, made by a Detective Constable would match Rule 3 and be permitted by the SPoC:

Request 3: [POL.DOM.DETCST] requests [BEHAVIOUR] of [PERSON] with [DOM.INVST] from [SOC.REHAB] within [Start] to [End]

However, if the Detective Constable requested this information in relation to a vehicle parking offence (Context=VPO), as in Request 4, the request would not match a defined rule and be denied by the SPoC.

Request 4: [POL.DOM.DETCST] requests [BEHAVIOUR] of [PERSON] with [VPO] from [SOC.REHAB] within [Start] to [End]

In this case, the SPoC may return the following message:

Vehicle Parking Offence is not a defined role in Information-Sharing Agreement.

3 Horizontal management of data protection and data sharing requirements

In the example discussed above, highly domain specific rules on data protection and data sharing were modelled together with information sharing agreements between individual agencies. Our claim is that the data protection aspect that is in this way hardwired into the information exchange protocols is not, as so often perceived to be, a hindrance to efficient police work. Rather, it is essential to allow offline trust relations between individuals in the various agencies to be replicated as "trust in the system" online. Different agencies do not lose ownership of the data that they have collected for their own purposes – in line with the data protection principle of purpose roundedness of data collection. Nor are they required to abide by a centrally imposed conceptualisation of their work. Rather, the negotiated settlements preserve the unique understanding of the various agencies, and can give them confidence that any information sharing that the system permits also aligns to their own understanding of the relevant data protection law. Medical workers, whose work is driven by a care ethos, can therefore more confidentially share information with police requesters without having to fear that the police's very different approach to data protection issues - as noted by PBR – overrides their own understanding of the relation they have with their patients.

Rules can be quickly updated as changes in the information sharing agreements, or changes in the covering legislation occur. However, this also means that a considerable degree of domain specific coding of legal rules has to take place. This raises amongst other questions the issue of this approach is scalable. We therefore continue our discussion of privacy enhancing technologies and inter agency sharing by looking at a more abstract approach to the problem that nonetheless preserves significant features of the system described above, and remains true to its empirical and motivational underpinning. In the terminology of PBR, the resulting system allows a global, or horizontal, management of the conflict between data protection and data sharing requirements

This part of the paper outlines the safi.re (Structured Analysis, Filtering and Integrated Rules Engine) framework[1] which creates a formal structure for the abstraction, governance and implementation of trust relationships and security policies. Trust and security become the more abstract, and explicitly defined, counterparts of what we saw above in terms of highly domain specific firewall rules. Once these have generated a trust relation between the parties that is based in the knowledge that any data sharing between them will abide by mutually acceptable rules, we can forget the content of these rules (which run in the background) and focus instead on ensuring that only trusted parties can access the information. It can be used as a full end-to-end solution for policy abstraction, implementation and controlled access to services, or can integrate each of the elements *as a Service* to existing applications.

The safi.re architecture has been used in a number of projects including with health and social care, including with the TSB-funded DACAR project with Chelsea and Westminster Hospital in London which focused on creating an e-Health Cloud within a hospital environment (Fan et al 2007). This used a novel method of defining the ownership of the data, and providing rights infrastructure for the citizen (or patient) to define the rights of access to their data. This work has since been extended within a number of projects including the TSB Trusted Service project, which has focused on integrating both digital and human trust, to provide a fully integrated and holistic care infrastructure, which integrates primary and secondary health care with assisted living while preserving patient privacy. (Fan et al 2012; Ekonomou et al 2011; Lo et al 2012).

As discussed in the previous section, another important area for information sharing is within the holistic care, where information from different public sector agencies can be used to improve the care of citizens. This might relate to sharing information on a child for concerns posted within health, social care, education and policing, where concerns within just one of these domains would not be seen as a major concern, but when aggregated across several of these, it might result in the concerns being escalated to the point where an action plan is initiated. The work has thus evolved into projects which involve information sharing for Child Protection, which involve the increasingly typical a multi-agency approach (see e,g, Blyth 1990; Reder and Duncan 2003). As there is information held within each of the public sector agencies, it is important that accesses are well managed and controlled for the rights for the access to data.

3.1 Key questions within a trust framework and terms

A simple abstraction of access is that an **accessor** from a specific domain accesses a service within another domain. This access can be in terms of their identity and/or role, and which might involve some form of relationship to the subject of the access. For example a GP might access the health record of one of the patients, where they

[1] http://safi.re

have the role of being the actual GP of the patient, given them higher access rights than if they were just a GP.

Overall the key questions in the design of a trust framework include:

- **How is the infrastructure segmented into domains, and what is the structure of these domains?** A key part of the safi.re architecture is that small sub-domains can exist within a larger domain definition. This allows for micro-policies to be defined within the larger domains, and for macro-ones to exist for the connection between domains.
- **Who is trusted within the infrastructure, and how do they map to the defined domains?**
- **Who are the data owners of the core data, and who will be trusted to govern this data?**
- **What services do these domains offer, and what is the formal definition of the services?**
- How are users identified and what attributes they need to consume a **service from another domain?**
- **How do we handle delegation, consent, and relationship, where the accessor is accessing information related to a subject, and has some connection to them?**
- **How are anonimization and sanatization dealt with across cross-domain access?**
- **How do we wrap the data access up into well-managed service access points?**

3.2 Safi.re Architecture

In modern service-oriented infrastructures a user must gather claims to consume a service. Too often the service is bound to a specific authentication infrastructure which limits the scalability of the provision of the service. For more dynamic infrastructures there is no direct communication between the service and the gathering of the claims around identity and the attributes required to consume a service, . It is the focus of the Trust and Governance infrastructure to define a contract which binds these terms of service together. This contract pre-defines the requirements for the claims to the service, and then is trusted to actually issue the contract for the user to consume the service. The Safi.re architecture abstracts the trust relations from well-defined policies. A trusted broker will then pass the requirements for a user to consume a service, and the Trust Framework will provide back the claims that are required to be able to consume the service. The user will then gather the claims, and the broker then passes these to the Governance Engine for it to check its running rules for rights to the service. If these are acceptable it will issue a service token to consume the service, which can be given back to the user, via the broker (or the service can be invoked on their behalf, and the link to the service can be returned to the user).

As with our approach outlined in 3 above, a key element of the Trust Framework is the concept of role, relationship, consent and delegation, where an role can claim rights of access to a referrer. In this way the owner of the data can have rights of access based on their role (such as whether they are a GP), their relationship (such as whether they are the GP of a specific patient), their consent (whether someone has given them rights), or their delegation (where they have given delegation of authority to another person).

3.3 Trust and Governance as a Service (TaaS and GaaS)

With the complex relationships that organisations have in rights of access to services, it is becoming increasingly more important to abstract and fully define the trust and the levels of access to services. Safi.re thus provides the ability to extract the trust relationship between two domains, and then implement this as a set of rules. These are thus defined in the Trust Framework and the Governance Engine, which can be easily integrated into existing applications. Figure 3 outlines a basic use case, where a broker deals with the requests form a user. It will then use the Trust Framework to define the requirements of the claim to a service, and the Governance Engine to check these rights against the actual rules of access to a service. Dynamic trust relationships can be built up for identity and attribute providers, and how these map to the role, relationship, consent or delegation that an individual has to consume a service. The service itself can be involved by the broker or a service token can be sent back to the user for them to give to the service. In this way both legacy services and new trusted services can be integrated into the infrastructure.

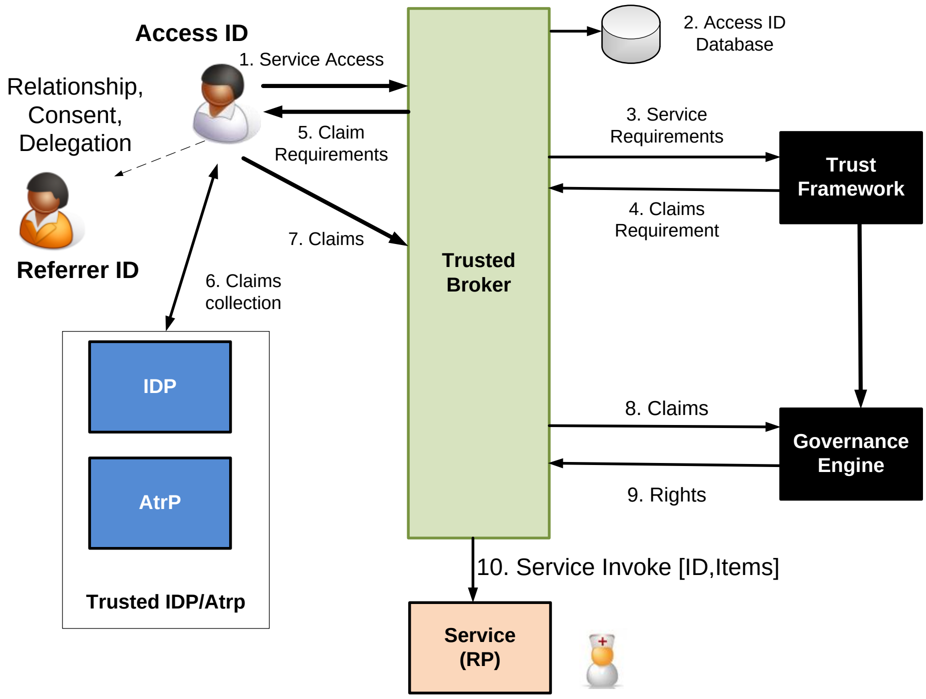

Figure 3: Trust and Governance as a Service

3.4 Safi.re Gateway Engine

Safi.re can also implement a filtering gateway which takes the rules from the Governance Engine, and runs them with a Gateway Engine, which then directly runs the rules, in a similar way that a network firewall will implement the filtering of network packets. Figure 4 outlines the full integration where the abstraction of the trust relationships, then created the rules, which are then implemented within a gateway, which will control access to the services based on the trust relationships. This type of architecture fully implements an end-to-end solution for trust relationships, where the requirements can be audited and reviewed, with control of each stage. It also can integrate with a wide range of stakeholders, using trusted identity infrastructures.

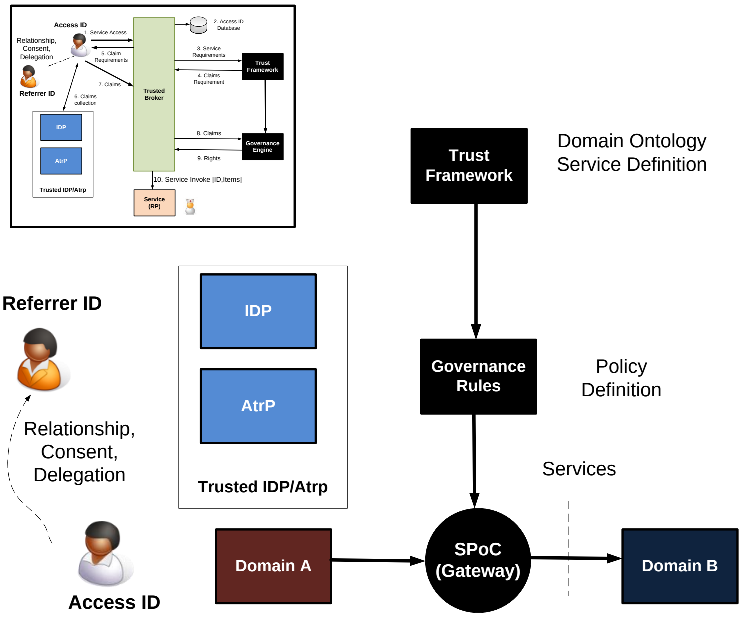

Figure 4: Full integration

3.5 Novel Binary Decision Diagrams for Modelling

What we need in addition to the formalism described above in 3 is the integration of a formal trust framework and implemented rules, and then modelling of complex trust relationship between domains using a patent pending method of Binary Decision Diagrams (BDDs).[2] BDDs are rooted, directed, acyclic graphs originally proposed by Lee (1959) and Akers (1978) in 1978 to graphically represent Boolean functions. BDDs originate from binary decision trees which are rooted, directed trees that can be used

[2] US Patent Application No 13/739074

to represent Boolean functions. For example, the decision tree illustrated in Figure 5 represents the Boolean function f(x; y) = (x ∨ y).

The concept behind this form of representation is that each non-terminal node (circle) in the decision tree denotes a variable. In the example illustrated in Figure 6, the variables are x and y. The node refers to a test of the variable's binary value, 0 or 1, with the edges of the node representing the paths taken by either possible value. The path represented by the dashed (low) edge corresponds to the case where the variable is assigned a 0, and the path represented by the solid (high) edge corresponds to the case where the variable is assigned a 1. The bottom (square) terminal-nodes of the tree represent the Boolean constants 0 and 1. Hence, the value of any Boolean function may be evaluated for any given number of variables by starting at the root (top) of the tree and following the path at each node, as determined from the value of the variable that the node represents. This process is repeated until a terminal-node (bottom) is reached. The value of the Boolean function, either a 0 or a 1, is represented by the value of the terminal node.

A difficulty with representing Boolean functions with decision trees is that if the function contains a large number of variables, then the decision tree representing that function will also be very large. Figure 6, for example, represents the binary decision tree for the function f(x; y; z) = (x ∨ y ∨ z). A comparison of Figure 5, which represents a Boolean function with two variables, x and y, with Figure 6, which represents a Boolean function with three variables, x, y, and z, illustrates that there is an exponential relationship between the number of variables in a the function and the number of nodes in the decision tree which represents that function. With increasing numbers of variables, therefore, the size of the decision trees used to represent functions increases exponentially. As can be expected, the decision trees of complex Boolean functions can quickly become very large and difficult to use.

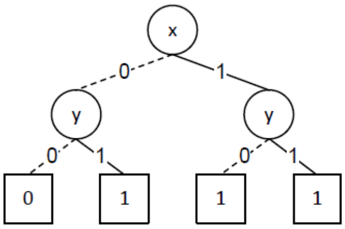

Figure 5: Binary decision diagram for the function f(x; y) = (x ∨ y)

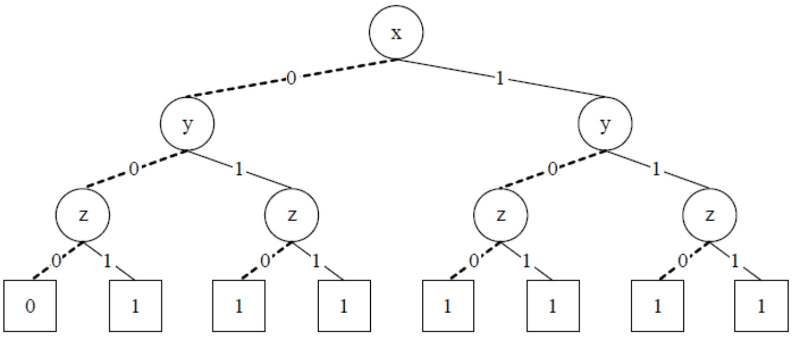

Figure 6: Binary decision diagram for the function f(x; y; z) = (x ∨ y ∨ z).

3.6 Reduced Ordered Binary Decision Diagrams (ROBDDs)

In 1986, Randal Bryant proposed a solution to this problem in (Bryant 1986) by introducing algorithms for reducing binary trees and ordering the variables in a function. The process of reduction consists of merging any isomorphic sub-graphs for the decision tree. Any parent node which has child-nodes that are isomorphic is considered redundant and is removed. Applying this process to the decision tree for the Boolean function f(x; y) = (x ∨ y), as illustrated in Figure 7, it is evident that if the first node, x, is 1, then the value of the second node, y, has no effect on the terminal node value of the Boolean function: whether y is 0 or 1, the value of the terminal nodes is 1. This means that the where node x is 1, child-nodes of y are isomorphic. Node y can then be considered redundant here and removed. The result is the reduced decision tree illustrated in Figure 7. Similarly, applying the reduction process to the decision tree for the Boolean function f(x; y; z) = (x ∨ y ∨ z), illustrated in Figure 7, yields the reduced decision tree shown in Figure 8. Reduced decision trees allow a much more compact representation of Boolean expressions than non-reduced decision trees.

Bryant also highlighted that the size of a decision tree for a given function is dependent on the ordering of the variables in that decision tree. For example, the decision tree for the Boolean function f(w; x; y; z) = (w ∧ x) ∨ (y ∧ z), given a variable ordering of w; x; y; z, is illustrated on the left-hand diagram in Figure 8.

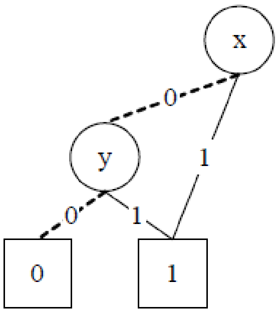

Figure 6: Reduced Binary Decision Diagram for the function f(x; y) = (x ∨ y).

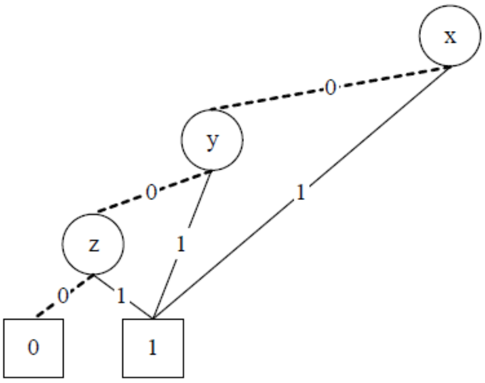

Figure 7: Reduced Binary Decision Diagram for the function f(x; y; z) = (x ∨ y ∨ z)

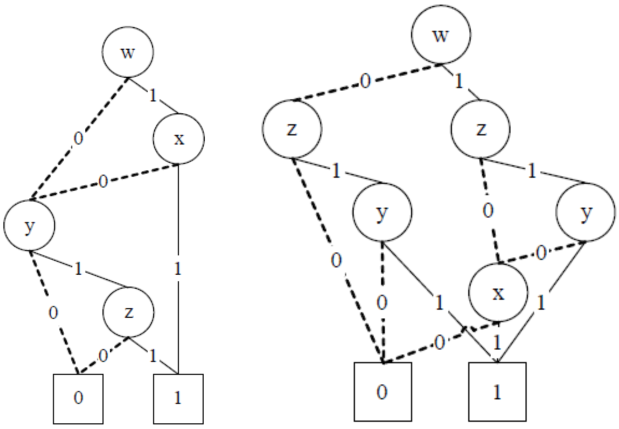

Figure 8: Reduced Binary Decision Diagram for the function f(w; x; y; z) = (w ∧ x) ∨ (y ∧ z) with variable ordering of w; z; y; x.

If the variable ordering for the same function was now changed to w; z; y; x, the resultant decision tree will be more complicated, as illustrated in the right hand side of Figure 8. Hence, an optimal variable ordering will produce the simplest, and therefore smallest, decision tree for a given function, while sub-optimal orderings will produce larger and more complex decision trees for the same function. However, as shown by Bollig and Wegener (Bollig and Wegener1996)], determining the optimal variable ordering for a Boolean function is an NP-complete problem that often requires trial and error or expert knowledge of domain-specific ordering strategies.

Decision trees which have been reduced and ordered are referred to as Reduced Ordered Binary Decision Diagrams (ROBDDs), or commonly shortened to just Binary Decision Diagrams (BDDs). A key property of the reduction and ordering restrictions introduced by Bryant is that the resulting BDDs are canonical (Bryant 1992). This means that the BDD for any Boolean function, for a defined variable ordering, will always be isomorphic. This property has made BDDs ideal for use in formal equivalence checking. In the electronic design automation process, for example, BDDs are frequently used to formally prove that two circuit design representations exhibit the same behaviour.

3.7 BDDs in Policy Modelling

A novelty of this proposal is to exploit the unique properties of Binary Decision Diagrams (BDD) to model complex sets of policies, in a form that is readily machine-executable, and to extend these to the information-sharing domain. The work of Hazelhurst et al. (1998) with firewalls identified key constituent fields in access-list rules and translated these into bit vectors representing BDD variables. This research applies a similar methodology to information-sharing where a set of information-sharing policies can be modelled as a decision diagram, once a specific variable ordering scheme has been selected. The modelling of a set of policies as a BDD provides a number of significant advantages, including providing an efficient lookup mechanism for an information-sharing request as well as providing a graphical representation of the overall policy set. As rule sets become larger and more complex, they become difficult to interpret and maintain (Hazelhurst 2000). Modification of the rule set, by either adding new rules or removing existing ones, or even changing the order of rules has a significant impact on the behaviour of the policy-based system. As noted above this is a key requirement for modelling rapidly changing legal environments Hence, analysis and validation of large, complex rule sets is essential in ensuring that high-level directives are enforced.

3.8 Domain Modelling using BDDs

The core of the patent is the linkage with the trust framework and the governance rules. In order to simply the access to data from domains, the method exposes only well-managed services to define the trust relationship. Within this the model defines a number of modelling elements, including:

- **Permission**. This is a simple permit or deny for access to a service.
- **Domain**. This relates to the domain that an accessor is contained within, and is used to create the holder to the domain ontology.
- **Organisation**. This relates to an organisation with a given domain.
- **Unit**. This relates to a unit with an organisation.
- **Role**. This defines the role that an accessor has in access a service within another domain.
- **Relationship**. This defines the relationship that the accessor has to the data being accessed.
- Action. This defines a CRUD (Create, Read, Update or Delete) access to a service and its associated data.
- **Attribute**. This defines an attribute of the object to be access, such as for a health record.
- **Object**. This defines the actual access target, such as for a specific person.
- **Context**. This defines the content of the investigation (which can be used to define certain risk levels for access privilege escalation.
- **Compliance**. This defines the audit/compliance reasons for the access.

The trust framework then defines the usage of each of these fields, and rules are written which implements them. A sample rule is thus:

> [Permit] [Police.Police_Force_A.*.Sergeant] with [*] relationship [R] [Unique_Identifier] of [Child] with [Abuse_Investigation] context from [Social_Care.Child_Protection_Agency_B.Records_Unit.Records_Admin] with Compliance [Human_Rights_Act_1998]

Overall the BDD model uses a binary representation for each of these fields, and which builds-up a rule definition with the binary representation of each of the possibilities for the fields. For example if there are four roles, we can represent them with:

- 00 – Constable
- 01 – Sargent
- 10 –Superintendent
- 11 – Chief Superintendent

These rules then use the BDD to determine if there are issues within in the governance rules related to:

- **Redundancy**. This is where one set of rules is already included within the trust rules already defined.
- **Shadowing**. This is where a rule is higher up in the set of rules, and matches all the conditions that match in the current rule, such that the shadowed rule will never be activated.
- **Generalisation**. With this a rule is generalisation of another preceding rule if it matches all the packets of the preceding rule.
- **Correlation**. Two rules are correlated if the first rule in order matches some of the fields of the condition of the second rule and the second rule matches some of the fields of the condition of the first rule.

The team have mapped out the modelling of these and have submitted a patent on the basis of this, and now require to scale-up the implementation to cope with real-life trust relationship. Full details at contained in the patent application [16].

3.9 Simple Example

We can now develop further the example discussed previously in 3, a request for highly sensitive data in a child protection case. This example describes in detail the steps needed to translate a set of information-sharing policies to a BDD. List 1 shows a sample list of policies. A '*' or 'Any' is used to denote redundant fields, or redundant portions of fields. Redundant fields are not translated into binary as they represent variables that are not evaluated by the BDD and, hence, do not form part of the Boolean function. Where an entire field is redundant, it is entirely excluded from the binary representation and where only a portion of a field is redundant, only the relevant portion is translated while the redundant portions are shown using 'Xs'.

Listing 1:

Policy 1: This policy <permits> <ANY> requester, with <ANY> relation in <ANY> context, to request to <read> a <child's> <Health History Record> from the <Records Admin> of the <Records Unit> of <Child Protection Agency 'B'> under compliance of the <Data Protection Act>

Policy 1:

Compliance (DPA)	: 1
Requester (Any)	: not checked by BDD
Relation (Any)	: not checked by BDD
Context (Any)	: not checked by BDD
Object (Child)	: 1
Attribute (Health History Record)	: 01
Owner (SocCare.CPA-B.RecUnit.RecAdmin)	: SocCare : 10 : CPA-B: 10 : RecUnit : 10 : RecAdmin : 10
Action (Permit)	: 1

The Boolean function corresponding to Policy1, ignoring redundant fields, is a logical conjunction of all of the above fields in the format shown in Listing 2. Listing 2 represents Policy1 expressed logically as an 'if-then' conditional statement.

Listing 2:

Permit: Compliance ^ Owner ^ Object ^ Attribute

Listing 3: Rule1 expressed as an if-then conditional statement.

```
if        (Compliance = 1) ^
          (Owner = 10101010) ^
          (Object = 1) ^
          (Attribute = 01),
then (Action = Permit)
```

4 Conclusions

This paper tried to bring computer science and socio-legal theory into a dialogue. We began by a socio-legal analysis of the problem government agencies in the UK face when trying to reconcile data protection rules with overarching data sharing requirements. We used in particular the influential study by Perri, Bellamy and Raab to develop a conceptual understanding of the nature of the tension. Form this we gained a number of key distinction: Vertical, domain specific approaches to manage the tension between privacy and data sharing can be distinguished from horizontal, or global approaches. Situations where the citizen volunteers data for his own benefit can be distinguished from situations where the data acquisition and data sharing is in the public good, with only indirect benefits for the individual. These different categories result in praxis in a plethora of different approaches and legal frameworks to manage data protection while ensuring efficient delivery of government services.

PBR conclude that in the majority of cases, this can result in marginalising privacy interest, contradicting the stated ideal of privacy and efficient service deliver reinforcing each other. The question our paper asked of the computer scientists then became a quest for formal approaches to privacy compliant data sharing that "squares the circle" and resolves, to the extent that this is possible, the apparent conflict between the two.

A key concept that emerged as part of both our empirical work with key stakeholders in pubic secretor agencies and through the literature was the notion of "trust": only if I can trust that my data tis sufficiently protected by robust privacy policies will I volunteer it as a citizen, only if I can trust the partner agency to use the information to they get from me as far as possible in a way that protects the privacy of my clients and sources, and acknowledges my professional understanding of my role and the integrity if my work will I be willing to share my data with them. Data Protection preserving ICT data sharing infrastructures and trust preserving ICT infrastructures therefore go hand in hand.

In Section 3 of this paper, we put these ideas into practice, by developing an information sharing syntax that directly embeds key ideas form data protection law into the data exchange protocols between agencies. By being privacy compliant by design, this approach should foster trust between agencies with vastly different professional attitudes and self-understandings – differences that as PBR have shown also influence the way they think about privacy. By allowing them to formulate their respective roles, concepts and understandings, and by providing an abstract formal concept of "criminal investigation context", we mimicked the balancing act between privacy and public interest prescribed by law, while acknowledging local differences and differences in professional attitudes.

The same concept of choice, consent and determination over one's data was then used in the final part of this part, where we abstracted further from the specific situation of individual agencies in a specific data sharing context, to a more abstract generalised approach that preserves privacy in cross-agency sharing through the proxy of "trust" – I trust only those partners with information I holed where the abstraction of the relevant privacy and data protection rules gives me reason to trust them. The exchange between legal and political theory and informatics that this paper attempted then shows at the very least how on the conceptual level, the issue raised by Perri, Bellamy and Raab can be answered through a computational, trust enhancing and privacy compliant information sharing infrastructure. For the computer scientists, this means that the alignment of their thinking with key socio-legal concepts gives us good reasons to believe that an approach that so far has been developed for and tried in a hospital setting only, can indeed be a seen as a model for privacy compliant data exchange across government departments in general .

5 References

Lee, C. Y. (1959). Representation of switching circuits by binary-decision programs. Bell System Technical Journal, 38(4), 985-999.

Akers, S. B. (1978) "Binary decision diagrams," IEEE Transactions on Computers, vol. C-27, no. 6, p. 509,.

Bellamy, C. "Joining-Up Government in the UK: Towards Public Services for an Information Age." *Australian Journal of Public Administration* 58.3 (1999): 89-96.

Bench-Capon, T. "Neural networks and open texture." *Proceedings of the 4th international conference on Artificial intelligence and law*. ACM, 1993

Beynon-Davies, P. (2007). Personal identity management and electronic government: The case of the national identity card in the UK. *Journal of Enterprise Information Management*, *20*(3), 244-270.

Blyth E, Milner J. 1990. The process of inter-agency work. In Taking Child Abuse Seriously: Contemporary Issues in Child Protection Theory and Practice, The Violence Against Children Study Group. Unwin Hyman: London.

Boer, A., Van Engers, T., & Winkels, R. (2010). Traceability and change in legal requirements engineering. In *AI Approaches to the Complexity of Legal Systems. Complex Systems, the Semantic Web, Ontologies, Argumentation, and Dialogue* (pp. 74-92). Springer Berlin Heidelberg.

Bollig, B and I.Wegener, "Improving the variable ordering of obdds is np-complete," IEEE Trans. Comput., vol. 45, pp. 993–1002, September 1996.

Bratley, Paul, et al. "Coping with change." *Proceedings of the 3rd international conference on Artificial intelligence and law*. ACM, 1991.

Bryant, R "Graph-based algorithms for boolean function manipulation," Computers, IEEE Transactions on, vol. C-35, no. 8, pp. 677–691, August 1986.

Bryant, R. E. “Symbolic boolean manipulation with ordered binary-decision diagrams,” ACM Comput. Surv., vol. 24, pp. 293–318, September 1992.

Choudrie, Jyoti, Vishanth Weerakkody, and Stephen Jones. "Realising e-government in the UK: rural and urban challenges." Journal of Enterprise Information Management 18.5 (2005): 568-585

Ekonomou, E., Fan, L., Buchanan, W., & Thuemmler, C. (2011, November). An Integrated Cloud-based Healthcare Infrastructure. In Cloud Computing Technology and Science (CloudCom), 2011 IEEE Third International Conference on (pp. 532-536). IEEE.

Fan, L et al. "SPoC: Protecting Patient Privacy for e-Health Services in the Cloud." eTEL-EMED 2012, The Fourth International Conference on eHealth, Telemedicine, and Social Medicine. 2012.

Fan, L W Buchanan, C Thuemmler, O Lo, A Khedim, O Uthmani, A Lawson, D Bell, DACAR Platform for eHealth Services Cloud, Cloud Computing (CLOUD), 2011 IEEE International Conference on, 219-226

Francis, B., Soothill, K., & Dittrich, R. (2001). A new approach for ranking ‘serious’ offences. The use of paired-comparisons methodology. The British Journal of Criminology, 41 (4),726–737

Hazelhurst, S “Algorithms for analysing firewall and router access lists,” University of the Witwatersrand, Johannesburg, South Africa, Tech. Rep. TR-WitsCS -1999-5, 2000.

Hazelhurst, S, A. Fatti, and A. Henwood, “Binary decision diagram representations of firewall and router access lists,” University of the Witwatersrand, Johannesburg, South Africa, Tech. Rep. TR-Wits-CS-1998-3, 1998.

Lee, C. Y. (1959). Representation of switching circuits by binary-decision programs. Bell System Technical Journal, 38(4), 985-999.

Lo, O., Fan, L., Buchanan, W. J., & Thuemmler, C. (2012). Technical evaluation of an e-health platform. IADIS E-Health.

Perri 6, Bellamy, C., & Raab, C. (2005). Joined-up government and privacy in the United Kingdom: managing tensions between data protection and social policy. Part II. *Public Administration, 83*(2), 393-415.

Raab, C., & Bellamy, C. (2005). Joined-up government and privacy in the United Kingdom: managing tensions between data protection and social policy. Part I. *Public Administration*, *83*(1), 111-133.

Rauhofer J '”Privacy is dead, get over it!” Information privacy and the dream of a risk-free society' (2008) *Information and Communications Technology Law* Vol. 17 Issue 3, pp. 185-197

Reder, P., & Duncan, S. (2003). Understanding communication in child protection networks. *Child Abuse Review*, *12*(2), 82-100.

Schafer, B. (2011). All changed, changed utterly?. *Datenschutz und Datensicherheit-DuD*, *35*(9), 634-638.

Stranieri, Andrew, et al. "A hybrid rule–neural approach for the automation of legal reasoning in the discretionary domain of family law in Australia." *Artificial Intelligence and Law* 7.2-3 (1999): 153-183.

US Patent Application No 13/739074, The Court of Edinburgh Napier University, Short Title: Binary Decision Diagrams, IP Title: Improved Information

Uthmani, O., Buchanan, W., Lawson, A., Scott, R., Schafer, B., Fan, L., & Uthmani, S. (2011). Crime risk evaluation within information sharing between the police and community partners. *Information & Communications Technology Law*, *20*(2), 57-81.