Cambridge Centre
for Housing &
Planning Research

Encouraging inter-regulator data sharing: the perceptions of regulators

A report for the Better Regulation Delivery Office,
Department for Business, Innovation and Skills

Paul Sanderson
Daniel Banks
Simon Deakin
Chihiro Udagawa

January 2015

UNIVERSITY OF
CAMBRIDGE

Centre for
Business Research

This research was commissioned by the Better Regulation Delivery Office (BRDO) of the UK Department for Business, Innovation and Skills (BIS) and carried out by a team from the University of Cambridge led by Paul Sanderson. The authors wish to express their thanks to BIS and to the numerous interviewees for giving their time to discuss the issues around sharing business data between regulatory organisations.

Dr Paul Sanderson
Daniel Banks
Prof. Simon Deakin
Chihiro Udagawa

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This report was commissioned by the Better Regulation Delivery Office (BRDO) as part of a programme of work aimed at promoting greater sharing of business data between regulatory organisations. The research on which this report is based investigated the perceptions of regulators on the subject – i.e. their views on what promotes or inhibits data sharing. Findings and recommendations are summarised below:

- There is a distinct gulf between large data-rich and small data-poor regulatory organisations. The former have the power to acquire the data they need and tend to question whether the benefits warrant the costs involved in sharing data with others.

- Inequalities between the data-rich and data-poor could potentially be reduced if a codified duty and power for regulatory organisations to share data was introduced - incorporating ideally the comply-or-explain principle, to allow for operational flexibility and to avoid the need for legislation to address existing prohibitions.

- Successful inter-regulator data sharing relationships are typically built upon pre-existing social and organisational networks. Regulatory organisations should consider where they have such resources and how these could be utilised to increase data-sharing where potential social and economic benefits can be shown to exist.

- Type of data is critical. Technical and legal considerations make data-mining of shared datasets of even simple factual information more resource-intensive than sharing discrete intelligence on criminal noncompliance - the latter benefits from Data Protection Act exemptions and the well-established networks of regulatory intelligence staff.

- Constraints on data sharing arise in part from the operation of the legal and regulatory regime governing data protection. Problems include uncertainty, complexity, and disproportionality of sanctions. A particular issue is the way the Data Protection Act overlaps with, or sometimes cuts across, statutes applying to specific public bodies.

- A perception that the sanctions available under the Data Protection Act are disproportionate to the objectives of the legislation appears to be having a chilling effect on data sharing between public bodies.

- A ‘National Data Strategy’, recommended in the Shakespeare Review, could help break down technical barriers to data sharing focusing perhaps on encouraging joint working on data sourcing and mandatory publication of data catalogues, as well as consideration of a new public sector data sharing standard.

- Some government departments already take a lead in supporting data-sharing amongst their reporting organisations. This is helpful as realising the potential benefits from increased data sharing requires such support, particularly to overcome technology issues which disproportionately affect smaller, data-poor organisations.

1. INTRODUCTION

1.1 Aims and objectives

This report is the result of research commissioned by the Better Regulation Delivery Office (BRDO) as part of a programme of work aimed at promoting greater data sharing between regulatory organisations. BRDO and others have already commissioned work to highlight the value of information sharing[1] and centralised data collection,[2] the real and perceived legal barriers to data sharing,[3] and a case study on data sharing across an entire sector.[4] As far as possible, data should be shared, both internally and externally, with other regulators. Indeed, the Regulators' Code for non-economic regulators includes a duty to ensure requests to businesses for regulatory data should follow the principle of "collect once, use many times."[5]

However, the work to date has not examined in detail the perceptions of regulators themselves – the people on the ground who make both strategic and day-to-day decisions about data-sharing. Why do they opt to share or not share data, and in what circumstances, with what data? What do they consider are the factors inhibiting or promoting inter-regulator data sharing?

The research underpinning this report thus focused specifically on regulators' responses to the research questions, which were to:

(i) Determine the types and importance
(ii) of regulatory data most readily shared between regulators;
(iii) Investigate the following three aspects of data sharing:[6]
  a. Organisational[7]
  b. Legal
  c. Technical
(iv) Consider where the potential costs and benefits of increased data-sharing may be realised;
(v) Outline where support and intervention could usefully promote greater sharing between regulators of business information;

The research method and findings are presented below, followed by three short expert commentaries on the organisational, legal and technical issues raised. The report concludes with a number of recommendations on the ways in which regulators could be further supported to meet the objective of increasing levels of data- sharing.

[1] e.g. Greenstreet Berman Ltd (2013) *Research Results: What is the Value in Regulators Sharing Information?*, Birmingham: Better Regulation Delivery Office.
[2] Local Better Regulation Office (2011) *Data Collections: Report to the Welsh Regulators' Forum*, Birmingham: Local Better Regulation Office.
[3] Law Commission (2013) *Data Sharing Between Public Bodies: A Consultation Paper*, Consultation Paper No 214, London: Law Commission..
[4] Independent Farming Regulation Task Force (2011) *Report: Striking a balance: reducing burdens; increasing responsibility; earning recognition*, Department for Environment, Food & Rural Affairs.
[5] Better Regulation Delivery Office (2014b) *Regulators' Code*, Birmingham: Department for Business, Innovation and Skills.
[6] Analogous to the Centre of Excellence for Information Sharing three pillars: technology, information governance and culture (http://informationsharing.co.uk/altogether-now/).
[7] Originally culture, but organisation is somewhat broader. It includes structure as well as the reproducible values, norms and practices etc. that constitute organisational culture.

1.2 Research method

The data, findings and recommendations in this report are distilled from a series of 42 meetings held in the summer of 2014. Sixty regulators, two members of trade bodies and one external expert participated. All had some knowledge of their organisation's activities around inter-regulator data sharing. Together they represented 30 UK regulatory organisations or organisations with a strong interest in regulation and, by extension, data sharing between regulators. Some were known to BRDO or the researchers, others not. Interviewee organisations were selected from previous contacts and from a list of UK regulators but in order to get a spread of different characteristics in terms of size, sector, data use etc., three sectors were especially targeted: air travel, food, and small retail. Respectively these sectors can be characterised as having a dominant single sectoral regulator, several key co-regulators, or numerous regulators and inspectorates but no dominant single regulator. The full list is appended together with the typical roles of the interviewees. The majority of participants were regulation directors and managers, front-line regulators or inspectors with responsibilities for licensing, compliance or enforcement. A number of CEOs, as well as policy, legal and technical support staff also took part.

The primary objective of the meetings was to explore the perceptions of interviewees on what promotes or inhibits regulatory data-sharing. They were encouraged to speak openly of their experiences. The researchers assured the interviewees their comments would be, as far as possible, anonymised - consistent with typical social science practice for research of this type. However, none made this a condition of speaking and many stated they would be prepared if required to be quoted on the record.

Interviewees are identified in this report by a number and are classified solely by the extent to which they engaged with data-sharing and whether they considered their organisation to be data-rich or data-poor – a dichotomy that appeared early on in the interview phase of the work. Note that some of the interviews involved more than one interviewee. The use of the identifiers in the text indicates therefore that one or more interviewees made a substantive comment on the issue but only the first few to discuss the issue have been identified, for reasons of readability. It follows the number of interviewees attached to any issue is not indicative of relative frequency. A few issues discussed were either not clearly attributable to any interviewee or may have been raised by the researchers for consideration. These have been labelled PS.

To ensure data commensurability the researchers developed an interview guide (appended). The questions were based on a focused review of the academic and grey literature, but essentially the researchers let the interviewees speak for themselves, guiding the conversation around areas the interviewees themselves considered of greatest importance. Thus the guide was used to check issues had been discussed rather than to administer survey-style questions.

To trigger the discussions interviewees were asked about:

(i) Their experience of inter-regulator data sharing and the extent to which their organisation engaged in such activity;
(ii) Legal issues around data sharing, and in particular, the Data Protection Act, and;
(iii) Technical issues around data types, storage and access.

The interviews were recorded and summaries including quotes were made. The recordings of the interviews have been analysed, sorted into topics and then into the three categories: organisational, legal and technical – data/IT. There is inevitably a fair amount of overlap. For example some preliminary recommendations are included within the findings as they arose during discussions of, and are inseparable from, current practice.

1.3 Data-sharing focused literature

A focused review of the organisational, legal and IT/data literature informing data sharing in and between public organisations was undertaken. The primary objective was to generate questions and issues for discussion in the interviews.

The evidence from the organisational behaviour literature on the interplay between the inter-organisational, intra-organisational and interpersonal, in determining levels of public sector information sharing was especially helpful.[8] Following an essentially classical organisational approach, Yang and Maxwell (2011), writing on information sharing in public bodies, suggest the extent to which an information is shared is informed by:

(i) External relations and environmental factors such as law and politics;
(ii) Internal organisational factors such as size, overall resource levels, the type of tasks undertaken, levels of trust, culture[9] and
(iii) Interpersonal relations, values, experiences and interactions of relevant staff

For example, on the former, the extent to which the relevant government department treats its reporting agencies as dependent or independent can be critical. On the latter, how a civil servant understands public service can determine responses to open data vs privacy issues.

However, it is not a simple predicable relationship. Writing on e-government inter-operability, Pardo et al. (2011) suggest there are 16 factors that affect capability to share across organisational boundaries grouped around how the organisation approaches sharing in its governance, operational and strategic management, and its information policy as well as its technological readiness and existing levels of cross-boundary cooperation.[10] Clearly, successful and sustained sharing built on long term relationships will probably require some form of *quid pro quo,* a point made by Bradford (2011) in a report on sharing data to combat fraud.[11]

Knowledge around the law and ethics on data sharing was also instructive. The lack of any general power for public authorities to share data is in part offset by quite wide data sharing powers either conferred directly by enabling or other statute, or implied by their functions.[12] In the UK the Data Protection Act (1998) is the main legislation covering sharing of personal information although public authorities also need to consider Human Rights, the law of confidence and any prohibiting statutes. According to the Information Commissioner's Office (ICO) guidance, "Personal Information' is anything that can identify a living individual."[13] This definition is not usually affected by the context of those data – i.e. personal information held by a public authority or in a business database is still personal information for the purposes

[8] Yang, T.-M. and Maxwell, T. A. (2011) 'Information-sharing in public organizations: A literature review of interpersonal, intra-organizational and inter-organizational success factors', *Government Information Quarterly,* 28(2), 164-175.

[9] On which see previous work for the Local Better Regulation Office: Staite, C. (2011) *Assessment of Regulatory Culture: A Literature Review undertaken for LBRO*, University of Birmingham: Institute of Local Government Studies.

[10] Pardo, T. A., Nam, T. and Burke, G. B. (2011) 'E-government interoperability: Interaction of policy, management, and technology dimensions', *Social Science Computer Review*.

[11] Bradford, M. (2011) *The challenges and opportunities for sharing data to combat fraud*, Newark: Regulatory Strategies Ltd.

[12] As construed over time by the courts – e.g. Attorney General vs Great Eastern Railway Company (1880) or through the super gateway provided by Schedule 15 of the Enterprise Act (2002).

[13] Information Commissioner's Office (2012b) 'Determining what is personal data', available: https://ico.org.uk/media/for-organisations/documents/1554/determining-what-is-personal-data.pdf [accessed 30 December 2014].

of the Act which makes the position of data on sole traders problematic. Treating sole trader data as business rather than personal could – depending on the precise circumstances - compromise both the intentions of the DPA and Article 8 of the European Convention on Human Rights. The balance between privacy and access to data is pivotal. This dichotomy epitomises the difference between Kantian and Utilitarian ethics. The former demands that no one be treated as a means to an end (the right to privacy), while the latter entails creating the greatest happiness for the most people (effective e-government).

The uncertainty civil servants face in determining how to strike a balance contributes to sub-optimal levels of data sharing.[14] Like these incompatible philosophies, it may be that the right to privacy and the concerns of e-government can never be entirely reconciled – which means those making decisions on data sharing will always face a risk. Anonymisation of personal data reduces those risks but can never eliminate them. As the Law Commission (2014 7) put it:

> *"There is always a risk that the combination of information held about an anonymised data subject may enable them to be identified or that it may be possible to re-identify the individual by combining other data with the anonymised data*" (Batista and Cornock 2009)[15]

The variations in the way data are collected, formatted, and stored also have significant impacts on data sharing.[16] While regulators may use fairly limited forms of data, mostly factual, some will be subjective, particularly intelligence stored as free text comment. However, the main issues around data sharing identified in the literature are to do with how to understand data collected by an external organisation. The most commonly raised issues are the use of inconsistent terminology or jargon and inconsistency of measurement units or categorical grades as well as the inclusion of unnecessary information for data re-users (Bache et al. 2013, Ruusalepp 2008).[17,18] Other issues include problems around extraction where shareable data are intertwined with un-shareable data (Akers and Doty 2013)[19] and relatedly, the loss of information in reformatting (Bache et al. 2013).

Data storage and security are also concerns. Ruusalepp (2008) outlines three types of data storage designs in use in research communities - which are equally relevant to the regulatory context:

(i) Centralised data storage – each regulator transfers their datasets in a common format to a single location - i.e., data repository or national data bank;
(ii) Federated data storage – each regulator has a physically separate dataset but they are combined in a single virtual space to form a virtual common dataset;
(iii) Distributed data storage – each regulator has a physically and virtually separate dataset and transfers (part of) its data to other regulators upon request.

[14] Batista, L. and Cornock, M. (2009) 'Information sharing in e-government initiatives: Freedom of Information and Data Protection issues concerning local government', *Journal of Information, Law & Technology*, available: http://www2.warwick.ac.uk/fac/soc/law/elj/jilt/2009_2/bc/ [accessed 18 October 2014].

[15] Law Commission (2014) *Data sharing between public bodies: a scoping report*, London: HMSO.

[16] Council on Governmental Relations (2012) *Access to, Sharing and Retention of Research Data: Rights and Responsibilities,* [online], available: [accessed 22 December 2014].

[17] Bache, R., Miles, S., Coker, B. and Taweel, A. (2013) 'Informative Provenance for Repurposed Data: A Case Study Using Clinical Research Data', *The International Journal of Digital Curation,* 8(2), 27-46.

[18] Ruusalepp, R. (2008) 'Infrastructure Planning and Data Curation. A Comparative Study of International Approaches to Enabling the Sharing of Research Data', available: www.jisc.ac.uk/media/documents/programmes/preservation/national_data_sharing_report_final.pdf [accessed 22 December 2014].

[19] Akers, K. G. and Doty, J. (2013) 'Disciplinary Differences in Faculty Research Data Management Practices and Perspectives', *The International Journal of Digital Curation,* 8(2), 5-26.

All three prompt issues of data integrity and security, and require considerable cyberinfrastructure, which raises the issue of who pays – with larger data holders being wary of encouraging free riders.

"Technological challenges exist, however, due to the dispersed and heterogeneous nature of these data. Standardization of methods and development of robust metadata can increase data access but are not sufficient. Reproducibility of analyses is also important, and executable workflows are addressing this issue by capturing data provenance. Sociological challenges, including inadequate rewards for sharing data, must also be resolved" (Reichman et al. 2011).[20]

The answer is of course to design processes that clearly set out to address these problems. Of particular relevance here, Hulstijn et al. (2011) present a method for processing business data that is:

> *"… driven by quality management and process redesign approaches to deal with the unique characteristics of compliance reporting: legal data requirements, provenance, process compliance and multiple stakeholders. In particular, PPM[21] stresses strict adherence to an iterative development schedule, and shared conceptual models of processes, data definitions, technological infrastructure, and governance agreements."[22]*

Put simply they seek to avoid the internal silo effect, by for example, involving legal officers, operations, IT and data staff continually in the process of defining, formatting, processing, and using data, and in considering requests to share those data. To operationalise this they note that:

> *"In a political networked environment, project management methods like PPM need to balance between restricting participants' behavior and providing flexibility for participants to develop solutions which fit their needs."*

This emphasis on developing guidance on when to restrict and when to allow autonomous decision-making helps manage and limit any sense of uncertainty without taking an overly bureaucratic approach. In this way, it can be argued, confidence can be increased that decisions balancing, for example, data access and privacy or data security and data sharing, are correct, or at least defensible.

[20] Reichman, O. J., Jones, M. B. and Schildhauer, M. P. (2011) 'Challenges and opportunities of open data in ecology', *Science,* 331(6018), 703-5.

[21] Public Process Management.

[22] Hulstijn, J., Wijk, R. v., Winne, N. d., Bharosa, N., Janssen, M. and Tan, Y.-H. (2011) 'Public Process Management: a method for introducing Standard Business Reporting', in *12th Annual International Conference on Digital Government Research (DGo'2011)*, College park, MD, June 12-15, http://homepage.tudelft.nl/w98h5/Articles/ppm.pdf: http://homepage.tudelft.nl/w98h5/Articles/ppm.pdf,

2. FINDINGS[23]

Organisational and legal issues were named by many regulators as the key factors inhibiting data sharing (e.g. R9). Technical issues, although important and in need of addressing, were not seen by many as insurmountable - but only if government could be persuaded to fund the technology required to improve data-sharing for smaller organisations.

2.1 Organisational factors

Environment

The policy environment at both national and international level can affect levels of data sharing between public organisations. While encouraging more sharing of data in the UK may be desirable some interviewees claimed this should only really be done in concert with the EU and other international bodies if UK-based individuals and businesses are not to be comparatively disadvantaged. However, it was also pointed out that much could be learned from the good practice developed elsewhere, for example in Australia[24] and Canada[25], where data-matching across the public sector has improved the consistency of data held (R20, R21).

Changes in domestic policy and resource allocations also stimulate changes in the level of data sharing. As the regulated environment changes so regulators look for ways to acquire the necessary data or to acquire existing data from others at lower cost (R1, R15, R21). Ironically, sharing can itself be resource-intensive and at least one regulator noted that levels of sharing had dropped as a result of the abolition of some key partner data-sharing institutions. This led them to examine the cost as a whole of their data sharing arrangements, and to end some (R23).

Critical incidents can also drive increased data-sharing (R19) as regulators, and indeed businesses, seek to minimise the impact of such incidents and also to avoid reputational damage from any repetition. They do this by increasing resources allocated to data acquisition, agreeing more Memoranda of Understandings etc. (R13, R16) and because critical incidents often cross organisational boundaries, developing strategies for more effective inter-regulator cooperation (R39). The type of harm involved also can affect the resources regulators are prepared to allocate to data-sharing. Motivation to share both data and the cost of acquisition, is increased where acquiring more information can be identified as having the potential to prevent serious damage to the regulator, the public or the industry, for example in high profile health and social care incidents (R4).

Institutional location

Institutional arrangements affect all aspects of a regulatory organisation’s activities, not least data-sharing. While some sponsoring departments take a best-practice lead by encouraging data sharing amongst the regulators that report to them, perhaps centrally funding data-sharing schemes by hosting joint databases (R31, R32), others leave arrangements to their individual regulators, intervening only to arbitrate when disputes around data-sharing arise

[23] Note the data recorded in this section are primarily the perceptions of the regulators, as stated by them in the various interviews undertaken for this project.

[24] See http://www.oaic.gov.au/privacy/privacy-act/government-data-matching.

[25] Data matching protocols have been a component of Government of Canada policy for many years. See for example the various documents listed at http://recherche-search.gc.ca/rGs/s_r?st=s&s5bm3ts21rch=x&num=10&st1rt=0&langs=eng&cdn=canada&hq=&q=%22data+matching%22&wb-srch-sub=

(R34). An interviewee in the latter category reported that it could even be difficult on occasions to obtain data from their own department (R34). However, another interviewee commented they found government departments attitudes to data sharing “excellent compared to regulators” especially with respect to open data (R42).

Given the varied landscape of UK public administration, a regulator may also report directly to their Secretary of State, without any enabling legislation – a situation that seemed to provide one interviewee with a good deal of freedom of action and thus confidence in dealing with issues around limitations on sharing business data (R24). They had no limiting enabling legislation to deal with. There was though no consistency on distance from government influencing approaches to data. Some argued that arms-length bodies were less certain of their position so resorted to legal counsel more frequently (R32) while others suggested they tended to be keen to demonstrate their independence by not being seen to engage with other agencies (R33).

Resistance amongst regulators to sharing privileged information can also indicate the presence of a dominant profession amongst regulatees – they may mirror their regulatees’ lack of “appetite for transparency” (R23). Regulatory organisations are thus no different to other types of organisations. Their behaviour and culture reflects their experiences and origins, for example, whether they have developed regionally and retained a devolved open structure or were created purposefully as a focused, centralised authority (R25). Agency characteristics can also reflect the operating environment. A regulatory organisation that has responsibilities in a crowded sector composed of complex sub-sectors is unlikely to be able to regulate effectively without cooperating with others, making it a ‘natural collaborator’ for sharing data (R30).

Data-rich vs data-poor

The main divide amongst interviewees was between the larger data-rich organisations and the smaller data-poor. Large regulators generally have the power to demand all or almost all the data they require from their businesses (R2, R5, R36). For them there are few incentives to share. Indeed one interviewee remarked that beyond a joint competent authority arrangement they had no need to share data whatsoever (R36). They were able to demand all the data they required from businesses and on the few occasions where sharing was considered the decision was simple: “It’s either legal to share or illegal – we take legal advice and act accordingly - so what’s the problem” (R36). The easiest option for them when faced with data requests was to say “no,” citing their enabling legislation which, as with many regulators, limits data use to fulfilment of their remit (R10).

However there was one particular exception to this dichotomy. Some isolated inspectors in generally data-poor local authorities encountered few problems with sharing and, indeed, found the Data Protection Act and similar legislation empowering. This may be possibly because of their inspection role (compared to that of the rule-making regulator) and because they relied on maintaining communication at a distance with their counterparts in other areas and across professional boundaries in their own locality (R39, R40). They were thus more aware than most of the different cultures and rules governing behaviour in other organisations and were used to interpreting rules without constant referral to others, seeking legal advice only rarely.

Data-rich offenders?

A small number of regulators were identified by many (but not all) interviewees as “black holes” sucking in data but letting little out, one being labelled “a nightmare.”[26] The perceived default position of such organisations was simple: they turn down requests (R1, R38). Many however did acknowledge that those so identified generally held a large amount of confidential information about the public and were therefore more focused on privacy than their counterparts (R26). Moreover several interviewees identified cases where data were shared by what others deemed to be ‘nightmares,’ whether as part of a criminal investigation (R40) or as a paid service (R35). One even received a real-time data feed to aid joint criminal investigation activity (R30). Where such sharing did occur it was noticeable that it related almost exclusively to combating serious criminal activity where the data holder had a strong financial interest (R1, R2, R8, R24, R30). Ongoing discussions between BIS and large data holders were welcomed (R8), particularly if the complexity of some statutory gateways could also be addressed (R4).

Perhaps the benefits of data-rich organisations sharing with the data-poor are being under-sold by the latter to the former? As with any partnership it is important to highlight the *quid pro quo.* As one interviewee pointed out, access to corporate information on rogue traders operating at a national level could be beneficial to a number of regulators (R1). Another, more critical interviewee suggested data-rich organisations which do not share, for whatever reason, should be “named and shamed” (R15).

Resources

Resourcing data-sharing is difficult for smaller organisations. It is an expensive overhead as “the more you share, the more data are churned - so an ongoing relationship is required” (R14), especially to ensure compliance with the DPA. Several interviewees suggested government could facilitate sharing by providing specific resources for the purpose (R1, R4, R5). Certainly external support has been essential for getting pilot schemes such as Intelligent Regulatory Information System (IRIS - a discrete data-sharing partnership)[27] off the ground (R4), but such schemes also benefited from another essential input - the existence of champions within the partner organisations (R4, R31). In another example of (potential) support, one regulator of regulators is currently experimenting with pooling old data from a number of its reporting organisations to create a closed database in order to evaluate the compliance and enforcement benefits that could flow from operating a live real-time version (R25).

One way regulators can increase income for such purposes is to make data available on a commercial basis. A number of regulators supply data on a chargeable basis – and not just to the private sector. Providing data commercially, in a form suitable for exploitation by others, can be a source of income for data holders (R6). Of course, it can be argued this could lead to increased demands on businesses for more data, but if the income gained is used to offset regulatory fees, businesses could be encouraged to cooperate (R8). It should be noted that while some regulators do charge co-regulators for data-sharing, others do not (R11). The bases for these decisions are not clear and would require further investigation. Such inquiries could also usefully address how self-funding “trading fund”[28] regulators draw

[26] Unfortunately we were not granted interviews at these organisations.

[27] The pilot is in the testing phase involving the Health and Safety Executive, the Fire Service, Charnwood Borough Council and North West Leicestershire District Council.

[28] “A Trading Fund is granted standing authority to meet its outgoing expenditure from receipts and there is no detailed advance approval by Parliament of its gross income and expenditure. A Trading Fund framework is suitable in circumstances where agencies (or other parts of government) can charge for their goods or services through a genuine customer-supplier relationship and have a reliable income stream (at least 50% from

distinctions between commercial and regulatory activity (R24). It should be noted though that there was no evidence presented to indicate trading funds engaged in this activity to the detriment of regulatees or that it weakened their commitment to open data where feasible.

Networks

The pre-existence of networks of peers, arising perhaps from local partnership working, is important when establishing data sharing (R1, R4, R15). It is easier to establish trust amongst known individuals engaged in similar occupations at local level (R26) than amongst unknown individuals at national level. Relational distance is critical. The Better Regulation Delivery Office focus on piloting data-sharing at local level may well be the best strategy (PS). Indeed, it was suggested that national regulators could learn much from local sharing arrangements (R28).

Nonetheless, new networks can be created from scratch where sufficient motivation exists - for example where a new regulator is created or where an existing regulator lacks the data required to fulfil new tasks they have been given, or to mitigate budget cuts by replacing inspections with data analysis (R1). This requires appropriate venues for establishing networks to be made available (R23), for example, local partnership arrangements of various sorts, and is an area where government engagement could be beneficial.

Networks can also be concretised within existing data-sharing or intelligence-sharing organisations such as CIFAS (Credit Industry Fraud Avoidance System, R7) which can increase levels of trust (R23) and act as neutral third party data holders, able to access and make available to members, data some individual members would not be able to acquire alone (R26). Alternatively a regulator can internally set up an intelligence hub to act as a focus for shared data. This relies on the data being of sufficient value to external partners (R27). It was also suggested that the best partners were not necessarily those with the best data, but rather those with the greatest powers to obtain data (R34). On the negative side, a perception that government was being reorganised constantly was identified as damaging to the very networks that could be used to promote data-sharing. Indeed, dismantling such networks could cause levels of data sharing to fall (R22).

Confidentiality and trust

Businesses' trust in their regulators will vary but it was noted that some regulators rely on maintaining high trust levels with regulatees in order to obtain regular valuable intelligence on issues with serious consequences for consumer protection and safety. These trust levels could be compromised if businesses or whistleblowers within those businesses become concerned that confidential data will flow from the regulator to external organisations (R2, R14) and in a few cases the supply of data is in fact conditional on the regulator guaranteeing there will be no sharing (R3, R5).

Trust and cooperation may be enhanced however where data collection (via e.g. joint authority arrangements) lowers the cost to business (R11). In general, it was reported that businesses (excluding those closely involved in Primary Authority schemes) tend to be wary of greater data sharing amongst public bodies. Their main concerns were stated to be around (i) financial data, for obvious commercial reasons, and (ii) consumer-focused data, for both commercial reasons and because they fear such data may be misinterpreted, to their detriment (R18).

commercial activities)." Cabinet Office (2012) *Categories of Public Bodies: A Guide for Departments*, London: Cabinet Office.13-14

Regulatory task and role

The purpose of the regulatory organisation can affect attitudes to data-sharing. If a regulator is dealing predominantly with information that is factual and, by law, publicly available, then data-sharing may not really be an issue, although all regulators appear to hold some categories of personal data that are not made public and fall under the Data Protection Act restrictions (R6). Where the regulatory task is simply the supply by businesses of factual data then regulators may have little need to supplement those data externally.

In fact, regulatory role or task was only a major factor, and a positive one at that, when the data to be shared were specific intelligence about serious cases of noncompliance, rather than bulk datasets covering all regulatees. Interviewees identified as working on regulatory intelligence were more likely to share data. The issues of noncompliance with which they dealt were generally criminal in nature and thus covered by s29 of the Data Protection Act. Few complaints in respect of constraints on data-sharing were voiced by such interviewees. They stated generally they were able to satisfy any requirements for additional data using police powers and working cooperatively with relevant agencies. Some did however report frustrations around delays caused by DPA protocols and felt that, on occasions, such protocols had undermined criminal investigations (R6). Conversely, those working in civil rather than criminal law areas typically found agreeing data sharing with potential partners more difficult (R29).[29]

Unlike other interviewees, those involved in intelligence, perhaps working closely with the police, were also confident of their interpretation of legislation on data-sharing. Some seemed to adopt a form of 'just cause rationale' where those seeking data from others promoted a just cause, whether or not it was the primary reason for acquisition, in order to discourage a refusal from the data holder. It was noted by several interviewees that inter-regulator data sharing was easier to establish where the outcome could reasonably be said to fall within the various exemptions to the DPA; matters relating to criminal activity, terrorism and immigration (R4), (R15). Many intelligence departments within regulatory bodies employ former police officers because they have the experience and contacts needed to obtain and use the data required for criminal investigations (R38). Moreover, some reported mounting joint operations with the police, where the proximate reason given for sharing intelligence was whichever was considered most compliant with the DPA (R40).

Partnership working

While identifying commonality of interests (R21) or clusters of interdependencies (PS) could be a starting point for establishing inter-regulator data-sharing relationships (R20), setting up long term sharing arrangements from scratch, i.e. without the benefit of established networks, was considered potentially more problematic, or at the very least, resource-intensive. In such circumstances it was considered easier to reach agreement to share data for one-off projects or thematic studies of benefit to both parties (R2, R3, R4, R8, R23), and particularly where no statutory bars existed (R29). This would also make it easier to sell the benefits to other potential partners, so the initial design of multiple sharing arrangements is critical. (Although most data-sharing is bilateral or project based a few do allow limited access to their data-warehouses to groups of other, suitably secure, regulatory agencies with whom they have longstanding relationships, e.g. R11).

[29] On criminal vs civil law on these issues see Enterprise Act 2002 Part 9, sections 237-246, esp. 242:2 & 3; 245.

As a starting point, simple bilateral data-sharing is always easier to arrange than multilateral sharing (R8, R29). For example each organisation is likely at present to want its own legal advisors to check that any agreement is consistent with relevant legislation. While some of this could be reduced by the issuance of better guidelines, each organisation will still have to perform some checks against its own remit, enabling statutes etc.

Once a relationship is established a simple but powerful message from interviewees was that data-holders or indeed businesses can be encouraged to expand data-sharing activities if they receive positive and timely feedback from recipients on the uses of the data and outcomes (R19, R25).

Timeliness is important also. Data are becoming increasingly time-sensitive as we move to real-time collection and dissemination (R31). Providers can be reassured via feedback that the user understands the time and other limitations of the data and that the data are being used as intended and as allowed for under legislation, or consistent with any MoUs or data sharing agreements (R4, R8). The same can apply to businesses - the ultimate providers of data. Increased business-regulator data-sharing increases business-regulator contact and thus the opportunities for learning on both sides (R25), a valuable extra benefit given the post-Hampton[30] drive over the past decade to targeted, rather than periodic inspections.

Internal vs external data sharing

In most regulatory organisations, but especially the larger ones, the first point at which the legalities of data-sharing come into focus is internally, sharing between departments (R15). Where regulators carry out a number of discrete functions internal and external sharing may be subject to the same scrutiny, sometimes to the point where internal sharing becomes more difficult than external sharing, especially where internal sharing crosses professions with markedly different outlooks and approaches to privacy and data access (R39). Several interviewees commented on the way that internal silos reflected external silos in government as a whole, and indeed the wider economy.

It was considered that this silo effect could, to an extent, be addressed by government. Unless regulatory agencies are treated uniformly, with standardised formats for objectives, practices and roles, it is unreasonable to expect their approaches to one particular aspect of their work, data-sharing, to be harmonised (R21). Indeed, one interviewee described data sharing as “a mess, a patchwork quilt” with some regulators bound primarily by common law, others by statutory limitations (R26).

One regulator of regulators claimed the DPA did not affect their data-sharing activities at all, as they had maximal MoUs – allowing all data to be shared as they shared a common purpose – and were only dealing with other regulators who in turn had sharing agreements with their own regulatees (R24). In some ways they treated the regulatory organisations they oversaw as though they were located within their organisational boundary.

Process or outcome orientation

The extent to which individuals were process or outcome oriented seemed to affect attitudes to data-sharing. A focus on process was associated with a cautious approach to data-sharing in general, privileging privacy over access to information. A focus on outcomes produced the opposite effect. This was true within the same organisation or profession. Thus

[30] Hampton, P. (2004) *Reducing administrative burdens : effective inspection and enforcement,* London: HM Treasury.

interviewees with a protection focus dealing with process (R33), governance (R28) or professional ethics (R13) tended to take a more cautious, sometimes isolationist approach to data-sharing (R33) than colleagues concerned more with outcomes such as safeguarding the public (R1, R34). Outcome focused individuals might place little emphasis on privacy arguments (R4, R29).

Given this diversity of attitudes to data sharing, sometimes even within a single organisation, inter-regulator data-sharing could be encouraged further if there were dedicated secondments of staff closely involved with the collection, use, and in particular, sharing of data (R30). This would help participants to better understand the position held on data sharing by those in different organisations and professions. Secondments could focus on the way that professional values influence attitudes to data sharing, as well as organisational level issues such statutory bars, interpretations of the DPA in different contexts and so on. It would also be helpful if secondments involved both the data rich and the numerous smaller regulatory organisations that find it difficult to acquire data. This would help to address the sense that it is simply not worth approaching some departments as one can be “pretty sure they are going to say no” (R31).

2.2 Legal factors

Regulators, in common with most civil servants, are well aware of their constitutional position. Their freedom to act is limited by that position. They cannot act unless they have express or implied powers so to do (R16, R20). However, they exhibited a high degree of uncertainty about their powers to share data. This, combined with the risk they may be held personally liable for contraventions of the DPA, served to reinforce any organisational resistance and technical inhibitors to data sharing.

Data Protection Act

Interviewees were divided on the extent to which law hindered data-sharing but many expressed concern. The Data Protection Act was mentioned by very few as ‘enabling’ although s31 on Regulatory Activity was cited as helpful by some (R39, R40). Far more considered it inhibited sharing (e.g. R34), along with their own enabling legislation (R10). Even when the police make a data protection request to access information as part of an ongoing and time-sensitive investigation the process was described as slow and costly. For these reasons many interviewees considered the DPA needed revising. Some also considered the DPA, as currently drafted, failed to take account of recent technological advances (R9).

However, there was a sense that the DPA may be being blamed when other legislation inhibits sharing (R22) or when data collection agreements with businesses have been made unnecessarily restrictive by limiting data to a once-only use or for use in a particular contingency (R37) or simply where resources are scarce:

> *“I think people hide behind the legal thing as the reason they’re not doing [data-sharing] whereas in actual fact it is more of a cultural issue – i.e. Don’t really want to; more like hard work and exposes me to liability; our systems are [execrable] anyway and well actually the overriding thing is we haven’t got any money to spend” (R43).*

The Enterprise Act was also cited by some interviewees as causing sharing issues around what is and is not a regulatory authority (R27). However, most concerns were around perceived lack of clarity on the DPA and a perception that while sharing across the public sector ought to be the norm, it was not (R4). "Use any data of use" ought to be the starting point for ensuring compliance, it was claimed, and the DPA impeded this (R16).

While the Information Commissioners Office (ICO) views the Data Sharing Code[31] as the basis for a guide to best practice on data-sharing and claims its main concern is with the negligence of noncompliant organisations (R22), quite a few regulators consider the Act promotes a culture of not sharing, and called for, at a minimum, better guidelines (R38). Even where the ICO had indicated data might be legally shared, some cautious practitioners (R21), especially those working in Local Authorities (R14, R17) or for small data-poor organisations, were not prepared to take the chance unless they received confirmation from legal counsel on each and every individual case, which they could not afford. Given that individuals can in some circumstances be held personally liable under s55 of the Act their reluctance to take risks on data sharing is hardly surprising (R21, R27).

> *"If shared illegally an individual is responsible… comes down to this: if there was one legal opinion about how to share and which gateways to use, DPA interpretation etc. this would really help organisations"* (R27).

Some data holders expressed concern that sharing would expose the limitations of their data in respect of integrity and currency. For example a regulator may delete data after a period of time (R16) in line with their DPA data deletion policy (R22) but once shared those data may remain current elsewhere. Again, given the liabilities established under the DPA, data-sharing can be hindered, losing out in the calculation "balancing risk against trust" (R22). To combat the general level of uncertainty, some larger regulatory organisations have appointed specialist DPA representatives to each of their departments whose role, in addition to their core regulatory work, is to provide customised guidance on the meaning of the Act - and ensure compliance (R11). The development of a Competency Skills Framework was suggested for those working with the DPA, which may be of particular benefit to smaller data-poor regulatory organisations (R34).

Improving guidance on the DPA

Some interviewees called for a root and branch reform of the DPA but most would be satisfied with better, customised guidance (R22). This guidance should be sector specific, using sector language and exemplars (PS). This would enable regulators to far more easily assess the implications for protecting data in their own sphere, and importantly to understand the way the DPA impacts potential data sharing partners. While the law might enable data to be shared for "consistent purposes" based on "reasonable expectations" (R22) it is not so simple for regulators to distinguish how one use is consistent and another incompatible, and this uncertainty inhibits data sharing.

There was a sense that calls for increased data-sharing, particularly of personal data, could be unpopular with the public, and might cause regulators to be resistant (R20). In this context more guidance to the public could also be helpful, perhaps stimulating a debate around the appropriate balance between sharing data to promote consumer protection and protecting the privacy of individuals (R11) as well as more information on the amount of personal data held by the public sector compared to the private sector (R35).

Time and again interviewees accused the DPA of producing unnecessary uncertainty. As R42, said:

[31] Information Commissioner's Office (2011) *Data sharing code of practice*, Wilmslow: Information Commissioner's Office.

> *"I can't see it would be that difficult to have a very simple Q&A tool on a website that says, is your problem like this – like the old symptom checker thing - something similar like a relatively intelligent and transparent way of assessing the particular issue."*

The legal framework for sharing data at present is simply too complex. Some statutes allow sharing, others prevent sharing, some are far from clear, although to be fair, much of this pre-dates the DPA (R22).

> *"Statutory bars and DPA often cause concerns for regulators who are trying to share data. S55 is an individual offence. A key point is that you need guidance in order to share and need to know what you can and cannot do. [There is a] danger of being too fearful. It can be complex and require a detailed audit trail to feel comfortable. (R21).*

Data-sharing agreements, memoranda of understanding and statutory gateways

Most regulators engage in some form of formal data-sharing, many using multiple Memoranda of Understandings (MoUs), mostly within the UK (R21) although one or two share data only with supervisory bodies at EU or international level, the *quid pro quo* being the return of global regulatory data for domestic use (R3, R5). Very few stated they engage in absolutely no data-sharing. However, sharing arrangements can be costly to set up and time-consuming[32], sometimes to the point where regulators become frustrated and lose interest (R26), and statutory gateways are considered "far too complicated" (R4), politically sensitive (R20, R25), and restrictive due to over-specifying limitations on use (R26). They also require an understanding of the partner organisation (R21) which is easier where the partners already are members of a network (R23). Indeed, those who were not engaged in data sharing foresaw legal problems with data sharing agreements, either because they could be incompatible with existing agreements on confidentiality or because of statutory limitations.

MoUs themselves can be double-edged. While they enable sharing by setting out the processes by which data will be shared, they can, despite having no basis in statute, limit sharing, as in most cases sharing beyond the original stated purpose would almost certainly require a new MoU and further legal advice (R7, R13). To simplify matters, it could be helpful to follow the lead of the regulator of regulators that issues the same MoU to all the organisations it regulates. Although this is partly to ensure consistency in its treatment of the organisations it supervises (R24) a single template would be of particular assistance to regulators dealing with LAs and police forces where each authority seems to insist on negotiating individually (R38), despite the fact that most have the same legal status and purpose.

Once in place, much depends though on how the MoU is written and who is administering it on a day to day basis. They are after all, best practice mechanisms – separate from the legal powers necessary to share data – notwithstanding which, rigorous monitoring of MoUs can be resource-intensive (R34). On the other hand it was noted that where communication was sparse between a regulator's legal department and the data staff, the simple existence of a MoU was on occasions taken as authorisation to supply any data requested by the data sharing partner, without necessarily checking carefully the detail and importantly the purpose for which the data had been requested (R11).

[32] One scheme, conceived 14 months ago, has still not reached the first stage of an agreement on whether to share or not as each party has sought its own, inevitably conflicting, legal advice (R29) while another MoU took two years to put in place (R38).

MoUs, the DPA and enabling legislation can also separately or together prevent the re-use of data already shared, sometimes unreasonably so (R37). Despite having a legitimate case in principle for its re-use, an organisation may have to start the DPA compliance process again from scratch. This can even affect the internal re-use of data. R13 described a case where his organisation had realised some data they held were for restricted use so they had to delete their holding and repurchase at some expense exactly the same data again. This issue can be partially alleviated if overarching MoUs are issued outlining basic responsibilities of the parties and that the data can be used for *any* relevant statutory purposes. Individual MoUs to enable sharing of specific data in specific circumstances can then be issued in addition if deemed necessary (R21). A more far-reaching approach might be to consider greater re-use of data within any planed revision of the DPA.

Duty to share

While a national database was quite rightly rarely discussed as a practical objective (R16, R29) a statutory and/or codified "duty to data-share" for all regulators was suggested by some as a practical way to increase data-sharing (R1, R4, R8 etc.). The presumption should be that data are shared unless otherwise prohibited (R8). Perhaps data-holders should be required to explain where they have not complied with such a duty, following the comply-or-explain principle underpinning UK corporate governance (PS).[33] The idea was also greeted positively by many (e.g. R7) when raised by the research team.

Philosophically, this would in essence mean moving from the current focus on the DPA towards more of a focus on the FOIA (R8) as the standard. The FOIA is already used by some to assess sharing requests (R12) but noticeably, where this is the case, a duty to share was not considered particularly helpful (R11). There was a concern that it would require any current enabling legislation to be amended (R12) leading to considerable organisational resistance from larger regulators. If instated in the form of a code such fears may be groundless. Non-economic regulators already have to comply with the Regulators' Code.

A duty to share across the public sector would, it was argued, encourage standardisation of systems (R31). A counter argument is that it could lead to "a lack of innovation – if you rely on others for data, people become territorial" (R23, also R42). For example, the use of common definitions and fewer data collection points could provide a number of regulators with reasonably useful but sub-optimal data, which could compromise their effectiveness. Most interviewees however supported the idea.

Data catalogues

Most regulatory agencies need to work with others at some point to ensure they have the necessary data to carry out their statutory functions, particularly at the intelligence end of the data spectrum (R30). However, several interviewees posed the question, "do we know what we hold" and "do we know what is out there?" (R1, R8, R21). By way of response some also raised, and most were enthusiastic about, the idea of a duty to not only share data but, in order to facilitate this, a duty to indicate what data they held (in fairly basic terms) and the extent to which such data were likely to be shareable, and possibly the circumstances in

[33] See Seidl, D., Sanderson, P. and Roberts, J. (2013) 'Applying the 'comply-or-explain' principle: discursive legitimacy tactics with regard to codes of corporate governance', *Journal of Management & Governance,* 17(3). Also: Sanderson, P., Seidl, D. and Roberts, J. 'Disposition to comply with flexible regulation: regulatees' perceptions of the legitimacy of codes and the comply-or-explain principle', *Regulation & Governance,* (under revision, 2015).

which sharing would be positively considered (R5). This would be of greater benefit if it was applied across government, and beyond possibly (R1, R4).

Some interviewees pointed out that there was as far as they knew no single tool or database available that would enable them (or indeed, the businesses they regulated) to find out what was already made freely available by others (R8). Probably a greater issue, and one that is possibly easier to address, is the extent to which smaller organisations understand data. A duty to publish a data catalogue would mean having to categorise their data and its potential share-ability as well as explaining the value of the data (R16). The police National Intelligence Model may provide a suitable template and reportedly is used by up to two-thirds of those regulators dealing extensively with shared intelligence-based data (R21). This entails grading data to determine secure storage protocol, reliability, validity, uses etc. A catalogue of MoUs was also considered beneficial by some (R22) although larger regulators were not so keen, citing cost and lack of evidence of benefit to them (R36).

2.3 Technical factors

It can be argued (R13) that technological developments improving our capacity to collect and analyse data means increased data-sharing is almost inevitable, notwithstanding any perceived inhibitors (R15, R26). Regulatory data are no longer simply those data supplied by businesses but include data from other sources, including open access data. However;

> *"Increased data means increased noise, making it more difficult in some ways to discern uncompliant behaviour in data patterns. Overall though the rise in data leads to less variation in behaviours and slight improvement in behaviour"* (R13).

Defining data

Data can be considered to occupy a spectrum, with databases and the mining thereof at one end, and discrete pieces of intelligence consisting of one of more datum at the other. Moreover, definitions of data vary over time, so there is "no universal taxonomy across data sharers" (R23). This needs to be recognised in discussions of data-sharing and how to encourage it. Those who claimed to find data-sharing unproblematic referred almost universally to discrete pieces of intelligence on specific known or suspected cases of noncompliance rather than bulk data (R7, R21, R38). Resourcing intelligence sharing was not an issue for agencies actively engaged in intelligence gathering as a core part of their work.

On the other hand regulators dealing primarily with bulk data had little appetite for shared criminal intelligence other than on an occasional basis. Indeed, some interpreted legislation on sharing as enabling 'evidence' to be shared, but not 'intelligence' (R35). This appeared to be more of a definitional issue though. By intelligence, some interviewees meant specific data relating to particular cases whereas others used it to refer to the acquisition of general information on regulated businesses, from which cases of noncompliance might come to light - similar to data-mining. For many, the more compelling data-sharing driver was budget cuts - the need to acquire and use bulk data to target inspections in order to lower costs (R32). However, sharing bulk data is inevitably harder and more costly in terms of technology.

It is essential to understand the different sorts of data in order to place attitudes to data-sharing in context. One regulator claimed to pre-classify data held as high or mixed level, and system level, which indicated in outline form with whom those data could be shared

(R13). Closer inspection however showed that most sharing again related to discrete pieces of intelligence rather than bulk data, and that sharing was in fact fairly limited.

Types of data

Some types of data lend themselves to sharing more than others. Factual business data are often freely available elsewhere, albeit at a variety of different locations, so making such data feely available or shared in a useable form is hardly an issue for some (R6, R7, R8). On the other hand personal data about consumers and their children is far more sensitive and tends to instil a reluctance to share any data amongst concerned practitioners (R33), Free text data also concerns some as there is less input quality control so opinion may be conflated with fact (R38, R41). The reason given for acquiring data is also important to providers, primarily in respect of compliance with the DPA. Many data holders can supply for evidential purposes but not for intelligence or commercial reasons (R35).One interviewee commented that it was important to know what question to ask. Data, or even the underlying reason for requesting data, may be categorised under different headings in different organisations which means data may be under-utilised. It was argued that legislation could be helpful to make clear that “intelligence development is part of the investigation and enforcement process” (R35) and to ensure organisations take an organised approach to publishing lists of the data they hold.

Format

Format issues affect the transfer of data for both providers and receivers, internally and externally (R23), but in particular are an issue for resource-poor recipients who may experience problems affecting accessibility, manipulability etc. caused by software and browser incompatibilities (R17). Additionally it was pointed out that not all regulators can control the format in which their data are organised, for example, those providing data upwards to EU or international regulators that insist on prescribed formats (R2, R3).

Some interviewees suggested standardising data formats (R8) and ensuring that formats and interfaces are simple to use but as data collection moves to real time (R23), this may prove difficult to operationalise (R11). It can however be encouraged within families of regulators who already have high levels of trust and confidence in each other’s’ operations (R12) and those such as LAs who carry out the same functions in different locations. For example, there are several database systems in use within Trading Standards, on top of which, data may also have to be uploaded to several different external databases (R27). Unsurprisingly then, some regulators, particularly those with system responsibilities, considered IT incompatibilities to be the “killer” inhibitor for inter-regulator data-sharing (R14).

Categorisation of data

Typically regulators view data as occupying a continuum from (i) open access data, through (ii) limited access data provided under FOI requests and requests from other agencies on the basis of legislation, through to (iii) data protected by the DPA and other legislation perceived to be potentially prohibitive (R21). Alternatively (and probably incorrectly – R22) a number of regulators treated FOI requests as the yardstick for open access to data and thus categorised data as either freely available or not (R16). Whether correct or not this simpler approach was stated to cut the cost of data provision by reducing the number of specific requests requiring serious consideration (R16). Businesses, however, have concerns over how freely accessible data are understood by consumers (R19), as noted before, and by no means all data provided under FOI requests are public data or data that can automatically be shared with other regulators (R22).

Understanding data

The meaning of data is critical (R16). Even apparently straightforward data may require considerable explanation. Understanding another organisation's data may mean providing the recipient with a detailed explanation outlining sources, commensurability of those sources, statistical procedures used, limitations etc. (R2, R3). In this context the starting point for establishing successful data-sharing arrangements is for the provider to understand the capacity of the importing organisation to correctly interpret what they receive (R11) and to continually review the use of their data, not least because data definitions vary (R42), can be very political, and may be changed by end-users (R23).

One way to share data intelligibly is to aggregate it. Aggregation of business data may enable it to be used beyond the original purpose for collection (R15), perhaps to generate confidence in management ratings. However, such ratings are subjective in the sense of being based on businesses' responses to compliance with different sets of regulations. They need to be understood in context. Interviewees were more or less evenly divided on the extent to which data evidencing noncompliance in one area of a business's activities necessarily indicated the likelihood of noncompliance in other areas (R2, R5). For example, failure of a small business to file their annual accounts on time did not necessarily mean they took a cavalier approach to compliance in core business activities such as food supply (R15, R18). Even if there was a core business compliance failure it did not necessarily mean that the business was noncompliant in other regulated core areas (R43) except where the types of activity were similar (R38). Context is important. For example, noncompliance with food standards almost certainly represents a "higher risk for restaurants than newsagents" (R14).

Some took the opposite view on the benefits (or indeed pitfalls) of aggregation though, suggesting that, far from being aggregated, data should be left as raw as possible in order to leave space for professional judgement (R28) and to encourage the market to provide innovative applications (R31), to the benefit of both data providers and the public.

Data integrity

Data integrity problems, in the context of data sharing, affect the value of data shared. While some organisations have always attended to the integrity of their data, because of the nature of the risks with which they deal, others have paid it less attention, although most claimed to now take data integrity seriously (R9). In all cases sharing requires the provider to make an honest assessment of the integrity of their data and also the completeness of data held, explain data definitions, and if relevant, how they have changed over time (R16). Exposing data to others, by sharing or moving to open data, can feed existing organisational insecurities (R33) on the one hand or drive improvements on the other (R34) although ironically, the organisations considered most difficult to deal with on data sharing were also considered to hold data of the highest quality (R16). This is consistent with the notion that the best quality and most up to date data are those concerned with finance and earnings (R16, R38) or "for data integrity, follow the money."

Data sourcing

Diverse objectives such as the regulation of behaviour to reduce crime means gathering data from a variety of sources (R29), which can be costly. To reduce costs some regulators are out-sourcing direct compliance data collection and/or using commercial data (R11, R30) which lowers businesses' data supply costs. However, third party supply can inhibit data-sharing where the intellectual property rights are retained by the data supplier (R34) and/or where the provenance of the data is not clearly established (R35). Particular care needs to be taken where data are presented in court. Concern was expressed that uncertainties or inconsistencies in the legal basis for acquisition, retention and use of data may be used by defence counsel to weaken the prosecution case (R30). For this reason R22 suggested best data collection practice is to get express consent at point of collection.

Some of the organisational, legal and technical issues encountered in attempting to share data could be lessened if, prior to data collection, regulators consulted co-regulators and other regulatory stakeholders in their immediate environment, for example in order to agree data definitions and establish acceptable unique identifiers (R23, also R8, R12). On this, the EU requires regulatory data in a variety of formats leading to multiple requests to business (R12). It was argued by one interviewee that such inhibitors could be minimised at least on a domestic level if regulators used a single type of proprietary software rather than developing their own, potentially incompatible, bespoke alternatives (R31).

Data security

Data-sharing requires a secure environment for the transfer and storage of data and, where possible, data should be anonymised (R22) although there is of course the unavoidable trade-off between security and utility. For those dealing with more sensitive data, most sharing is of extracted secondary data (stored on standalone PCs) rather than local interrogation of networked raw primary data (R23).

On data transmission, not all regulators are on the GSi network, including some large agencies (R14) and thus are not able readily to satisfy others their data transmission security standards are comparable. It was reported that law enforcement agencies in general require secure email, the cost of which can be high (R27). Consideration could be given to standardising where feasible and assisting national and local agencies to meet security requirements (R1). Arguably, a major step forward could also be to produce a single technical and security data-sharing standard (R8) although, given the diversity of data and uses, this could be difficult to achieve. Regulators should however pay particular attention to data security because, according to the ICO, the major fines imposed on regulators to date have related to security infringements (R22). Also, data security breaches tend to lead to clampdowns on accessing and storing data that can impede data sharing (R34).

Hacking is also a potential problem. Factors such as user numbers and perceptions of the value of the data are relevant here. One way to combat hackling is to store more data on standalone PCs of course. Essentially, it has to be acknowledged that sharing requires trade-offs with security (R14) and security is another area where uncertainty and potential liabilities for contraventions inhibit data sharing.

3. COMMENTARIES

3.1 Organisational issues

Drawing on the wider organisation and information science literatures, Yang and Maxwell (2011) in their review of information-sharing in public organizations contend that levels of information sharing are best understood by examining the factors affecting inter-organisational sharing, intra-organisational sharing and interpersonal sharing.[34] In other words, and for our purposes, levels of data sharing between regulatory organisations are also informed by data sharing within those organisations and between members of those organisations – both within and across organisational boundaries.[35]

Interpersonal level data sharing

With regard to data sharing, the interpersonal level is concerned with the way factors such as an individual's values, experiences, capacities and networks affect their propensity to share data with others.

Process vs. outcome orientation

A number of interviewees displayed strong values that influenced their attitude towards data sharing, particularly in connection with protection of the vulnerable and sense of duty - a public service ethos, arising from membership of an autonomous profession or simply having a strong sense of vocation. The important element is what is valued, and whether it is intrinsic (outcomes for self) or extrinsic (outcomes for others) and how strongly it is valued. For example, an interviewee concerned with child protection outcomes saw little value in privacy or laws to uphold privacy. The need to protect a child from harm would always trump data protection concerns – a teleological argument (ends justifying means).

Recruitment in regulatory organisations is often from the sector being regulated which, in the case of regulation of the professions, may explain why professional values tend to permeate regulatory culture. Outcome oriented individuals in the same sector but in a different regulatory organisation displayed similar attitudes. Clusters of organisations grouped around strongly held shared values tend to share the way they articulate those values and indeed reinforce them by aligning their organisational systems and structures, and cultivating strong interpersonal relationships founded on shared understandings of what constitutes ethical behaviour.[36]

However, others working in the same sector, but in process-oriented or technical jobs, took the opposite view. They tended to take a deontological approach, focusing more on the moral principles underpinning the (privacy) obligations set out in the DPA, than outcomes. In

[34] Yang, T.-M. and Maxwell, T. A. (2011) 'Information-sharing in public organizations: A literature review of interpersonal, intra-organizational and inter-organizational success factors', *Government Information Quarterly,* 28(2), 164-175.

[35] While information is commonly said to be processed or interpreted data these terms are often used interchangeably - but context is important. The Information Commissioner's Office defines personal data as data which relate to a living individual who can be identified – which could be either raw data that could later be processed into identifying information or data already processed into identifying information - data and information are interchangeable. On the other hand the difference between bulk data and the processing of discrete extracts from bulk data to create information in the form of intelligence needs to be distinguished, as the addition of interpretation increases the value of the data, making it more likely to be shared.

[36] Cunningham, A. T., Bernabeo, E. C., Wolfson, D. B. and Lesser, C. S. (2011) 'Organisational strategies to cultivate professional values and behaviours', *BMJ Quality & Safety,* 20(4), 351-358.

doing so they are effectively conforming to the popular view that professions create protective organisational boundaries and limit information about noncompliance in the belief this maintains levels of public trust (although this would be an over-simplification).

Thus, in regulatory organisations particularly but not exclusively involved in regulating more ‘social’ type issues, there seemed to be a divide between a process and an outcome orientation. If a balance was struck within a regulatory organisation between privacy and data access it seemed to leave at least some individuals “optimally disgruntled,” as Wilson (1980 361) put it in respect of regulatory compromise.[37]

Buy-in to data sharing could be improved if all were engaged equally in the search to achieve a balance between privacy and access to data. One way to alter perceptions is to appropriate existing structures. Codes of behaviour can be useful in this regard. They are how a profession embeds and reproduces its values. Thus a codified duty to share data could have some effect - but not alone. Success depends on developing a strategy to keep the code visible and relevant, preferably with the support of regulatees (Frankel 1989).[38] Training is also important. Data sharing could for instance be included in a competency framework and within the Regulators’ Development Needs Analysis (RDNA) self-assessment tool.[39]

Secondments could also be useful - identifying regulatory organisations that are engaged in higher levels of data sharing and arranging secondments for staff involved or potentially involved in facilitating data sharing with organisations that are less engaged. Secondments are “a chance to uncover, question and find new ways of negotiating the unwritten rules that often govern various agency approaches“ (Cousins 2005).[40] “Our organisational and occupational rituals are second nature to us…as are the words we use… It is as if different languages are being spoken in this cross-cultural encounter and clear communication can become very difficult (Scott 1999).[41] Secondments also enable staff to ask questions about how data are understood and how decisions are being made on data sharing without the tensions that can arise in the course of inter-organisational negotiations to put a data sharing agreement or Memorandum of Understanding in place.

Networks and trust

Networks, and relatedly trust, are core components of interpersonal relations and central to establishing data sharing arrangements. Relational distance is critical with regard to networks. For example, contacts made via local partnerships and/or through sharing professional interests were mentioned by many of the interviewees. The networks they drew upon to share data were mainly informal but some were more formally constituted. Some were data repositories themselves, for example CIFAS (Credit Industry Fraud Avoidance System.)

In all cases trust is an essential component, especially in relation to potential personal liability for breaches of the DPA in the context of data sharing. Trust serves to reduce

[37] Anderson, E., Tritter, J. and Wilson, R. (2006) *Healthy Democracy: The future of involvement in health and social care*, London: INVOLVE and NHS Centre for Involvement.
[38] Frankel, M. S. (1989) 'Professional Codes: Why, How, and with What Impact?', *Journal of Business Ethics,* 8(2/3), 109-115.
[39] Better Regulation Delivery Office (2014a) *Common Approach to Competency for Regulators*, Birmingham: Better Regulation Delivery Office.
[40] Cousins, C. (2005) 'Short Term Placements: An Exercise in Organisational Culture Exchange', *Australian Journal of Public Administration,* 64(4), 81-89.
[41] Scott, D. (1999) 'Workshop on Interagency Collaboration for PANOC and DOCS staff, in C. Cousins' in [Article], Short Term Placements: An Exercise in Organisational Culture Exchange: Australian Journal of Public Administration 64: 4, 81-89.

uncertainty. It also reduces costs as it substitutes for expensive monitoring and compliance systems (Barney and Hesterly 1999 135).[42] Indeed, following the logic of transaction cost theory, high levels of trust not only reduce data sharing costs but also enable sharing to take place that would otherwise not occur as it would be too resource-intensive.[43] Creating trust is however difficult. Trust is generated over time as a result of frequent and successful interactions.[44] The literature offers no easy answers to building trust. The starting point therefore for increased data sharing has to be to build upon existing networks or interpersonal relations already established between those employed within regulatory organisations. Perhaps an obligation to clearly set out co-regulatory, joint authority and existing data sharing arrangements as well as future plans to increase sharing could be included within a codified duty for regulators to share data.

Intra-organisational level data sharing

A systematic approach to intra-organisational data sharing in regulatory organisations can serve as a best practice exemplar guiding external sharing. The extent to which data are shared internally is dependent on the interplay of a range of factors, including structure, culture, and task diversity. For example, while there has been a trend to encourage groups to share knowledge internally, consistent with the trend for organisations to become more agile and responsive to faster changes in their environment, there are still many features of organisational practice that conform to a more traditional bureaucratic model, where information flows are strictly controlled (Yang and Maxwell 2011 165).

In fact the level of data sharing activity within an organisation is the product of the interaction of a number of sets of factors. At the interpersonal level these include members' understandings of the costs and benefits of sharing and sense of ownership of the data. These are intermediated by organisational factors such as the organisation's capacity to deal with change, levels of IT, size, reward systems, internal politics, identity and image, and levels of trust, and at a meta-level, organisational structure and culture (2011 166).

Organisational structure

One of the more obvious influences on data sharing levels is arguably the extent to which decision-making is devolved within an organisation. Centralisation can inhibit internal communication and cooperative behaviours between divisions and departments.[45] One of the few regulatory organisations with almost no domestic data sharing suffered from an internal silo effect. It did not share data internally to any great extent.[46] The ability to share data across internal boundaries is also related to departmental function, whether the department is internal facing and process oriented or external facing and outcome oriented, as well as whether the relevant employees have external allegiances such as professional ethics.[47]

[42] Barney, J. B. and Hesterly, W. (1999) 'Organizational Economics: Understanding the Relationship between Organizations and Economic Analysis' in Clegg, S. R. and Hardy, C., eds., *Studying Organization*, London: Sage.

[43] Coase, R. H. (1937) 'The Nature of the Firm', *Economica,* 4(16), 386-405.
Williamson, O. E. (1979) 'Transaction-Cost Economics: The Governance of Contractual Relations', *Journal of Law and Economics,* 22(2), 233-261.

[44] Dasgupta, P. (1988) 'Trust as a Commodity' in Gambetta, D., ed. *Trust: Making and Breaking Cooperative Relations*, Blackwell, 49-72.

[45] Some however found no correlation at all. See for example Willem, A. and Buelens, M. (2009) 'Knowledge sharing in inter-unit cooperative episodes: The impact of organizational structure dimensions', *International Journal of Information Management,* 29(2), 151-160.

[46] As noted above, this was partly, but not wholly, a result of the different formats in which they were required to supply data to the international bodies to which they reported.

[47] See previous section.

Organisational culture

Schein (1992)[48] describes organisational culture as:

> *"A pattern of shared basic assumptions that a group has learned as it solved its problems of external adaptation and internal integration that has worked well enough to be considered valid and therefore, to be taught to new members as the correct way to perceive, think, and feel in relation to those problems."*

Organisational culture has been found to positively influence attitudes to data sharing within an organisation if the culture emphasises fairness, affiliation and innovation, and in particular when the norm is to share data (Bock et al. 2005)[49]. Inter-regulator data sharing may therefore be optimised if regulatory leaders establish data sharing as the norm, the default position. There are a number of tactics that can aid this process. A codified duty to share is one, but the trick with codes is indeed to establish the obligations contained therein as norms. As there are a number of sometimes confusing statutes affecting data sharing, and an air of uncertainty around what can be shared, any obligation to share needs to incorporate a degree of flexibility. This can be achieved by incorporating the concept of comply-or-explain into the obligation to share.[50] The norm then is to share data but where other laws take precedence to allow non-conformance - as long as a satisfactory explanation is given. In the corporate governance arena where this mechanism has been established for more than two decades, monitoring is normally carried out by the market – the investors. In the case of external regulatory data sharing the market would be other regulators who could no doubt highlight persistent refusals to share in appropriate venues. There is no particular reason why an internal market could not also help establish sharing as the norm. To an extent the mechanism is a form of naming and shaming!

Organisational resources

With the exception of trading funds, regulatory organisations are not set up explicitly to meet their outgoings primarily from the provision of goods and services. However, with public spending capped, many regulators do seek to charge for the provision of data in certain circumstances. It would be helpful to clarify the circumstances in which data sharing is or ought to be chargeable not least because sharing can be expensive, particularly for smaller regulatory organisations unable to proffer a *quid quo pro* to larger data providers. Smaller organisations also need continued support from BRDO and others to establish data sharing partnerships and indeed, on occasions to maintain existing levels of sharing, for example where institutional reforms disrupt existing networks and costly (time-consuming) rebuilding exercises need to be undertaken.

Of course, knowledge, information and data are also resources. Data produced by employees can be said to be property owned by the organisation (Constant et al. 1994)[51].

[48] Schein, E. H. (1992) *Organizational culture and leadership,* San Francisco: Jossey-Bass.

[49] Bock, G.-W., Zmud, R. W., Kim, Y.-G. and Lee, J.-N. (2005) 'Behavioral Intention Formation in Knowledge Sharing: Examining the Roles of Extrinsic Motivators, Social-Psychological Forces, and Organizational Climate', *MIS Quarterly,* 29(1), 87-111.

[50] See Seidl, D., Sanderson, P. and Roberts, J. (2013) 'Applying the 'comply-or-explain' principle: discursive legitimacy tactics with regard to codes of corporate governance', *Journal of Management & Governance,* 17(3). Also: Sanderson, P., Seidl, D. and Roberts, J. 'Disposition to comply with flexible regulation: regulatees' perceptions of the legitimacy of codes and the comply-or-explain principle', *Regulation & Governance,* (under revision, 2015).

[51] Constant, D., Kiesler, S. and Sproull, L. (1994) 'What's Mine Is Ours, or Is It? A Study of Attitudes about Information Sharing', *Information Systems Research,* 5(4), 400-421.

The extent to which data outputs are treated as corporate property determines attitudes to sharing that property (Jarvenpaa and Staples 2001)[52]. Clearly the predisposition to share data is likely to be higher where property rights are not an issue. The exception to this is where personal values outweigh property rights, for example where outcomes are privileged over process and/or where professional ethics override.

Inter-organisational level data sharing

Cross-boundary data sharing can entail complex interactions between participating regulatory organisations (Yang and Maxwell 2011 168). Organisations sharing data will be governed by different legislation and have different levels of technology and management styles. Origins, values and culture also play a part, as well as organisational function (Pardo et al. 2011)[53]. Some regulatory tasks are purely business oriented, hidden from public view, while others are highly visible and politically sensitive. This makes it difficult for data holders to know what data could be useful in another organisation and for data receivers to trust and fully understand the meaning of the data they receive. Trust between organisations depends on the length of time the relationship has been established, having the capacity to evaluate the trustworthiness of the counterparty, and the similarity or at least compatibility of organisational structure, cultures, norms and remit/legal powers etc. (see for example Rousseau et al. 1998).[54] Leadership has also been shown to be influential (Willem and Buelens 2009), an area where high- and intermediate-level workshops may raise awareness of the benefits of sharing and reduce uncertainty over the balance between privacy and data access.

Power imbalances

The core dichotomy that emerged from the research undertaken for this report was the very clear divide between the larger, data rich regulatory organisations and their smaller, data poor counterparts. The former had little need to acquire data and when they did, their power, arising from their size and data holdings, meant they were encounter little resistance, compared to smaller organisations. This power imbalance can be usefully explored using classic resource dependency theory (Pfeffer and Salancik 1978 1).[55] “To understand the behavior of an organization you must understand the context of that behavior—that is, the ecology of the organization.” In essence the theory points to the interdependencies within a population of organisations. These interdependencies breed uncertainty because the future actions of others are always to some extent unpredictable. Organisations act to reduce the uncertainties. Uncertainty is a source of power to those able best to withstand it. Thus smaller, data poor regulatory organisations face a double whammy from their environment. They face uncertainty about accessing sufficient data from their larger counterparts and are less able than them to withstand uncertainties over compliance with the DPA.

[52] Jarvenpaa, S. L. and Staples, D. S. (2001) 'Exploring Perceptions of Organizational Ownership of Information and Expertise', *Journal of Management Information Systems,* 18(1), 151-183.

[53] Pardo, T. A., Nam, T. and Burke, G. B. (2011) 'E-government interoperability: Interaction of policy, management, and technology dimensions', *Social Science Computer Review*.

[54] Rousseau, D. M., Sitkin, S. B., Burt, R. S. and Camerer, C. (1998) 'Not so different after all: A cross-discipline view of trust', *Academy of Management Review,* 23(3), 393-404.

[55] Pfeffer, J. and Salancik, G. R. (1978) *The external control of organizations: a resource dependence perspective,* New York: Harper & Row.

Relationship management

There are however strategic actions that smaller organisations can undertake, the most practical of which is to enter into alliances with other smaller organisations to, in the case of data sharing, improve their negotiating position on obtaining data from larger data holders (Provan et al. 1980).[56] They can also use network relationships to improve their position (Bae and Gargiulo 2004)[57], perhaps utilising network contacts in the government department to which they report. Finally there is impression management where data poor organisations simply develop strategies to highlight the joint dependencies, the *quid pro quo*, to data holders, creating "a vision of interdependence" (Ozcan and Eisenhardt 2009)[58] This is best illustrated by the interviewee working as isolated inspector in a local authority who solved the data access problem by joint working with the police and other powerful agencies, using inspection powers when police powers were insufficient for actions they wished to undertake.

3.2 Legal issues

The interview evidence reported above (section 2.2) suggests that the legal framework is a constraint on data sharing, at least to the extent of reinforcing organisational and particularly cultural factors. The problems associated with the legal framework are complexity, uncertainty, and a lack of proportionality of sanctions.

Complexity

The complexity of the law derives from the multiplicity of rules and regulations affecting the processing of data by public bodies. To some degree this is an unavoidable consequence of the constitutional position of public authorities. Statutory bodies can only act in pursuance of the powers conferred upon them by legislation. Many statutes contain specific provisions allowing data sharing, but it is not uncommon for more general but vaguely worded powers to be relied on to permit the processing of data. Government departments may be in a position to exercise common law or prerogative powers which go beyond the limits of those provided by statute, but the extent of any such additional powers is often particularly unclear. Where statutes provide for specific 'gateways' to data disclosure, the effect can be to clarify the extent of the relevant power, but in doing so the legislation may also confine it. In practice, it becomes necessary for civil servants and other public sector workers to have a detailed knowledge of the relevant law if they are to engage in data sharing.

The alternative to the present patchwork quilt of rules and regulations would be the enactment of a general rule allowing disclosure of data for purpose consistent with the exercise of public functions, or, less far-reaching, the widening of existing gateways on an individual sector basis. According to the Law Commission, conferring wider powers on public bodies will not necessary encourage more data sharing. This is because 'officials both lack confidence in the apparent scope of widely drawn powers and are wary of disclosing information otherwise than with the safeguards represented by the controls on onward disclosure that are typical of many of the existing gateway provisions'.[59] Our respondents, on the other hand, reacted positively, for the most part, to the idea of a duty to share across the

[56] Provan, K. G., Beyer, J. M. and Kruytbosch, C. (1980) 'Environmental Linkages and Power in Resource-Dependence Relations Between Organizations', *Administrative Science Quarterly,* 25(2), 200-225.

[57] Bae, J. and Gargiulo, M. (2004) 'Partner Substitutability, Alliance Network Structure, and Firm Profitability in the Telecommunications Industry', *The Academy of Management Journal,* 47(6), 843-859.

[58] Ozcan, P. and Eisenhardt, K. M. (2009) 'Origin of alliance portfolios: Entrepreneurs, network strategies, and firm performance', *Academy of Management Journal,* 52(2), 246-279.

[59] Law Commission (2014) *Data sharing between public bodies: a scoping report*, London: HMSO. at para. 11.48

public sector, which could be developed in the first instance through a code of practice. Our interview evidence also suggests that progress could be made by clarifying the terms of Memoranda of Understanding between public bodies and by regularly reviewing their content.[60]

Uncertainty

A major issue to emerge from our interviews is a perceived lack of clarity in the provisions of the Data Protection Act. The DPA cuts across existing statutory powers, so that even if data sharing is permitted or mandated by a specific statute, there must still be compliance with the DPA.[61] According to the Law Commission, many aspects of the DPA are not well understood, even by data protection practitioners.[62] The view of the Information Commissioner is that the DPA sets out a general framework, enabling the ICO to provide guidance which can be flexible and can evolve to meet changing needs.[63] Our interviewees, on the other hand, called for more sector-specific guidance as necessary to overcome uncertainty in the application of the DPA.[64]

As legislation in a relatively new field which is also affected by continuing technological change, the DPA is necessarily going to be subject to difficulties of interpretation. There is also limited scope to amend the DPA in the light of EU law requirements. However, the form of the DPA is not rigidly determined by EU law. As the Law Commission has pointed out, different Member States have implemented the Data Protection Directive in a variety of ways,[65] and there may be a case for reviewing the experience of other countries and benchmarking the DPA against regulatory best practice elsewhere. Pending a more radical revision of the DPA, our interviews suggest that there is scope for more precise guidance to be provided by the ICO on the issue of data sharing. Respondents' replies imply that the current Code of Practice on Data Sharing is too general to be of assistance on specific issues of the kind which arise in the context of information disclosure between public bodies.[66]

Lack of proportionality in sanctions for breach of data protection laws

Several of our respondents commented on the disproportionality of sanctions under the DPA and the chilling effect this could have on data sharing. The view of the ICO, as related to the Law Commission, is that these fears are generally unjustified. The Law Commission appears to share this view, referring to such fears as 'ill-founded'.[67] This is because individuals cannot be made the subject of monetary penalty orders under sections 55A-55E DPA, and because where monetary penalties have been issued by the Commissioner, this has so far been only for serious breaches of security.

It remains the case, however, that individuals may be subject in principle to a criminal prosecution for unlawful disclosure or receipt of data under section 55 DPA, which carries a maximum fine of £5,000 if the case is heard in the magistrates' court, and an unlimited fine if it is heard in a higher court. This offence is committed where the disclosure or receipt takes place without the consent of the data controller and requires the individual to act knowingly

[60] See above, section 2.2.
[61] Law Commission, Scoping Report, para. 1.56.
[62] Ibid., paras. 3.73-3.80.
[63] See section 2.2, above, and Law Commission, Scoping Report, at para. 1.100.
[64] See section 2.2, above.
[65] Ibid., at para. 3.33.
[66] See above, section 2.2.
[67] Law Commission, at para. 3.42.

or recklessly, making a prosecution of a public official acting in the course of their employment extremely unlikely. Even so, an individual whose negligence leads to a breach of the DPA by their organisation may be subject to disciplinary action or to dismissal.

The issue here is not whether fears of legal liability on the part of individuals or organisations are, objectively speaking, misplaced, but that they are genuinely held, which our evidence suggests is the case. If there is a misperception on the part of individuals of the scope of the DPA and of the possible consequences of a breach of its provisions, the problem may lie in the way in which the requirements of the Act have been communicated. It is also noteworthy that the sanctions provided for in the DPA are significantly more stringent than those set out in similar laws of other EU Member States.[68] Either way, there is a case for looking again at the proportionality of sanctions for breaches of the DPA that may occur in the context of data sharing.

3.3 Technical – IT/data issues

Although overlapping with the other main categories of 'Organisational' and 'Legal' addressed in this report, technical difficulties certainly add something new, and potentially expensive, to the mix. In this section various technical aspects are considered, many involving (directly or indirectly) the manifold requirements of the DPA on organisations handling personal information. In addition, some of the basics of data and IT management are addressed.

It may be worthwhile to consider all the following alongside the current low trust in public sector data security – albeit aggravated by sometimes disproportionate coverage in the media. Furthermore, the current 'culture of anxiety' (Law Commission 2014 107) is only accentuated by the growth of public-private partnerships.[69] Such trends arguably make the decision to share data even more arduous (Law Commission 2013 6), and continue the negative emphasis on privacy over efficiency.[70] Issues specifically around data security are addressed more fully below.

Data sourcing and collection

Inconsistencies in data collection and recording across public authorities can create difficulties for those who wish to share data (Office of the Children's Commissioner 2013 5).[71] This perspective was shared by a number of interviewees. Key factors included the purposes for which the data were originally collected, IP of any third-party data collection services and uncertainties over data definitions or collection methodologies employed. Clearly, where data are to be shared, it is advisable to provide details of this in any privacy notices or to make the stated purposes of data collection sufficiently wide to enable future flexibility.[72] In order to address many of these issues some interviewees suggested the use of a more collaborative data collection model.

[68] Law Commission, at para. 3.33.
[69] Law Commission (2014) *Data sharing between public bodies: a scoping report*, London: HMSO.
[70] Law Commission (2013) *Data Sharing Between Public Bodies: A Consultation Paper*, Consultation Paper No 214, London: Law Commission.
[71] Office of the Children's Commissioner (2013) *Office of the Children's Commissioner's response to the Law Commission consultation: Data sharing between public bodies,* [online], available: http://www.childrenscommissioner.gov.uk/content/publications/content_751 [accessed 22 December 2014].
[72] The DPA requires that new processing is not 'incompatible' to the stated purpose

Data storage and archiving

Inconsistencies in data storage and data handling across public authorities are also likely to create difficulties for those who wish to share data.[73] Ruusalepp (2008) presented three types of data storage designs in research communities, which are applicable to the current context:

(i) Centralised data storage – multiple regulators transfer their datasets in a common format to a single location - i.e. data repository or data bank
(ii) Federated data storage – each regulator has a physically separate dataset and information technology is used to provide a virtual common dataset at a virtual location
(iii) Distributed data storage – each regulator has a physically and virtually separate dataset and transfers (part of) its data to other regulators upon request.[74]

There are key differences of cost and utility in these three types of data storage designs. For example, centralised storage is likely to involve substantial investment while potentially providing the most comprehensive resource. Distributed storage, on the other hand, can be less expensive but more piecemeal. In any case, prospective sharers would need to reach agreement on the most appropriate storage design and on arrangements for backups, disaster recovery, auditing etc. Importantly, prospective sharers would also need to come to an agreement on the allocation of costs for the data sharing project. If costs at each stage of data sharing are placed on the principal, typically large regulator, then the typically smaller reuser of the data can be regarded as a free rider (Ruusalepp 2008, Berman 2008) – this is clearly to the detriment of those sharing regulators who must bear the expense of recording, formatting, storing, archiving and sharing data.[75]

For the purposes of storage and archiving, creation of metadata is of utmost importance to regulators. Metadata is descriptive information about data that explains the measured attributes, their names, units, precision, accuracy, data layout and ideally a great deal more. Most importantly, metadata includes the data linage that describes how the data were measured, acquired or computed (Gray et al. 2005).[76] A critical point in creating metadata is to make it computer-readable so that the contents of datasets can be understood by others (Akers and Doty 2013).[77] Metadata also need to include information on the data provider (Bache et al. 2013), i.e. the regulator or third party who collected and recorded the original information.[78]

All the above suggests that organisations may be most likely to opt for the less expensive, less complicated and more targeted distributed data storage design. Although easier to set up, this may not always be optimal to their needs.

[73] ibid.

[74] Ruusalepp, R. (2008) 'Infrastructure Planning and Data Curation. A Comparative Study of International Approaches to Enabling the Sharing of Research Data', available: www.jisc.ac.uk/media/documents/programmes/preservation/national_data_sharing_report_final.pdf [accessed 22 December 2014].

[75] Berman, F. (2008) *Got Data? A Guide to Data Preservation in the Information Age,* [online], available: www.sdsc.edu/about/director/pubs/communications200812-DataDeluge.pdf [accessed 03 January 2015].

[76] Gray, J., Liu, D. T., Nieto-Santisteban, M., Szalay, A., De-Witt, D. J. and Heber, G. (2005) 'Scientic Data Management in the Coming Decade', in *ACM SIGMOD Record*,

[77] Akers, K. G. and Doty, J. (2013) 'Disciplinary Differences in Faculty Research Data Management Practices and Perspectives', *The International Journal of Digital Curation,* 8(2), 5-26.

[78] Bache, R., Miles, S., Coker, B. and Taweel, A. Ibid.'Informative Provenance for Repurposed Data: A Case Study Using Clinical Research Data', 27-46.

Data security

According to the Law Commission, '…security issues…are…a key hindrance in data sharing and linkage' (Law Commission 2013: 7). Ongoing expansion of online services (Cabinet Office 2014a: 3) and the continuation of other decentralising trends – for example, social networks, Bring Your Own Device (BYOD) and cloud services – further compound these issues.[79] Given this expansion of 'attack surface', it is perhaps unsurprising that the cost to the UK economy of security breaches continues to grow (PwC 2014: 2).[80]

As is the case with many of the technical issues raised in this part of the report, proper data security requires continuous investment in infrastructure. This investment seems consistent with one of the overarching principles of the HMG Security Policy Framework:

> *Security must enable the business of government and should be framed to support HMG's objectives to work transparently and openly, and to deliver services efficiently and effectively, via digital services wherever appropriate* (Cabinet Office 2014b: 5).[81]

Indeed, the security of UK cyber space was listed as a 'Tier One' Priority Risk in the earlier UK National Security Strategy (Cabinet Office 2010: 27).[82] Where the necessary investment is not forthcoming, increased digital services, greater data sharing and other technological developments may be carried out to the detriment of data security.

Currently there are concerns that different public authorities use different, often incompatible, security standards. According to our interview evidence, resolving differences in email security, encryption standards, security classifications, retention periods etc. can be a challenge. Conversely, there are concerns from some commentators in the security industry over the development of a 'dangerous monoculture' (Lacey 2013); providing organised crime and terrorists with what could be regarded as a single point of failure for UK public authorities.[83]

The main driver behind data security in the UK is the Data Protection Act 1998 (DPA). In a recent publication the Information Commissioner's Office (ICO) usefully provided the following, typically broad overview of DPA requirements:

> *The Act applies to firms holding information about living individuals in electronic format and, in some cases, on paper. They must follow the eight data protection principles of good information handling. These say that personal information must be: fairly and lawfully processed; processed for specified purposes; adequate, relevant and not excessive; accurate and, where necessary, kept up to date; not kept for longer than is necessary; processed in line with the rights of the individual; kept secure; and not transferred to countries outside the European Economic Area unless the information is adequately protected.* (Information Commissioner's Office 2014: 1).[84]

[79] Cabinet Office (2014a) *Government Security Classifications*, London: Cabinet Office.

[80] PwC (2014) *Information Security Breaches Survey 2014,* Department of Business, Innovation and Skills [online], available: http://www.pwc.co.uk/assets/pdf/cyber-security-2014-technical-report.pdf [accessed 22 December 2014].

[81] Cabinet Office (2014b) *HMG Security Policy Framework*, London: Cabinet Office.

[82] Cabinet Office (2010) *A Strong Britain in an Age of Uncertainty: The National Security Strategy*, London: Cabinet Office.

[83] Lacey, D. (2013) 'Predictions for 2014', *David Lacey's IT Security Blog, http://www.computerweekly.com/blogs/david_lacey/2013/12/predictions_for_2014.html,* [Accessed 22nd December 2014].

[84] Information Commissioner's Office (2014) *Getting it right – a brief guide to data protection for small businesses, V3.1*, Wilmslow: Information Commissioner's Office.

Because the DPA only applies to ‘personal information’, it is vital that public authorities are able to identify this information to avoid expending resources on technical security controls unnecessarily. An example of current confusion in this regard involves the DPA status of information provided to regulators by sole traders.

Since 2010, ICO has been able to fine organisations up to £500,000 for serious breaches of DPA obligations. It has been using these powers with increasing frequency and, arguably, with a focus on public authorities. In recent times organisations fined by ICO have included Islington Borough Council, Aberdeen City Council, Ministry of Justice (on two occasions), North East Lincolnshire Council, Department of Justice Northern Ireland, British Pregnancy Advice Service (BPAS) and Kent Police. The largest reported fine was £200,000 for BPAS, a case that involved disclosure of sensitive personal information following hacking. The Ministry of Justice was fined almost as much, £180,000, for ongoing improper management of personal information in prisons. Such highly publicised failures in public sector data security create pressure on public authorities to increase vigilance on the personal information they hold (Cabinet Office 2014b: 3). They directly feed into the cycle of risk aversion and the current ‘culture of anxiety’ (Law Commission 2014: 107).

In practice, for larger organisations, the requirements of the DPA are addressed through ongoing certification (or effective alignment) to an enterprise data security standard, such as ISO 27001. Such certification provides an organisation with an Information Security Management System (ISMS) along with all the incumbent policies and procedures. Necessary security controls are determined by risk assessment – for example, more advanced security techniques might be required where information is high risk.

Where all appropriate measures are in place and evidenced, this may be sufficient proof for ICO that due diligence has been performed. Therefore, ICO seem unlikely to expect many organisations to successfully repel an Advanced Persistent Threat attack backed by a nation state. But, inevitably an ISMS will restrict what an organisation can do, or at minimum provide a list of, often costly, security measures that need to be in place before data can be shared.

The remainder of this section on data security looks at some particular issues public authorities seeking to share data may wish to consider. All of the following could have an impact on the final decision whether or not to go ahead with a data sharing project.

Anonymisation

According to ICO’s anonymisation code of practice, ‘anonymisation can allow us to make information derived from personal data available in a form that is rich and usable, whilst protecting individual data subjects’ (Information Commissioner's Office 2012a: 5).[85] In this way regulators may be able to protect privacy while taking advantage of shared datasets. However, anonymisation cannot be regarded as completely failsafe (Law Commission 2014: 27). In common with data aggregation, it can be a costly process that risks undermining the utility of the data (e.g.Oswald 2013: 8).[86] Therefore, the use of anonymisation needs to be carefully balanced to facilitate data release while maintaining the value of those data.

[85] Information Commissioner's Office (2012a) *Anonymisation: managing data protection risk code of practice*, Wilmslow: Information Commissioner's Office.

[86] Oswald, M. (2013) *Data Sharing between Public Bodies: Consultation Paper 214 – Response to Consultation,* [online], available: http://www.winchester.ac.uk/academicdepartments/Law/Centre%20for%20Information%20Rights/Publications/Documents/Law%20Commission%20Consultation%20No%20214%20Data%20Sharing%20between%20Public%20Bodies%20response%20from%20Marion%20Oswald.pdf [accessed 22 December 2014].

System monitoring

In the case where regulators are using shared servers, and in line with the requirements of ISO 27001, there is an obvious need to frequently, efficiently and securely verify that a storage server is faithfully storing the data (Ateniese et al. 2008).[87] This depends on adequate cyberinfrastrcture. Where data are shared more simply, each regulator will rely heavily on its own computers, external hard drives or internet-based storage. A mechanism of continuous monitoring of data storage facilities is required.[88] In the centralised case, the shared database should support investigations of misconduct, or misuse of stored data.[89] It should also provide a set of instructions on data storage and data retrieving as well assistive services.[90]

Encryption

Encryption is an important final layer of data security and all-important in this new era of distributed data systems. ICO has made it clear that it expects personal information to be protected by encryption[91] both when stored and when in transit.[92] In terms of data sharing, encryption may add overheads in the form of inter-organisation public key infrastructure management. Importantly, the organisations involved would need to ensure that their systems and encryption standards were compatible.

Data Classification

The Government Security Classifications replaced the previous policy on 2nd April 2014:

> *Government Departments and Agencies should apply this policy and ensure that consistent controls are implemented throughout their public sector delivery partners (i.e. NDPBs and Arms Length Bodies) and wider supply chain* (Cabinet Office 2014a: 3).

Such consistent classification across organisations would plainly be advantageous to data sharing initiatives. Any doubt over classification schemes used, or the details of implementation, might hinder efforts to share data.

[87] Ateniese, G., Pietro, R. D., Manchini, L. V. and Tsudik, G. (2008) 'Scalable and Efficient Provable Data Possession.', in *Proc. of SecureComm '08*, sprout.ics.uci.edu/pubs/a9-ateniese.pdf, 1-10.

[88] Hulstijn, J., Wijk, R. v., Winne, N. d., Bharosa, N., Janssen, M. and Tan, Y.-H. (2011) 'Public Process Management: a method for introducing Standard Business Reporting', in *12th Annual International Conference on Digital Government Research (DGo'2011)*, College park, MD, June 12-15, http://homepage.tudelft.nl/w98h5/Articles/ppm.pdf: http://homepage.tudelft.nl/w98h5/Articles/ppm.pdf,

[89] Akers, K. G. and Doty, J. (2013) 'Disciplinary Differences in Faculty Research Data Management Practices and Perspectives', *The International Journal of Digital Curation,* 8(2), 5-26.

[90] Ruusalepp, R. (2008) 'Infrastructure Planning and Data Curation. A Comparative Study of International Approaches to Enabling the Sharing of Research Data', available: www.jisc.ac.uk/media/documents/programmes/preservation/national_data_sharing_report_final.pdf [accessed 22 December 2014].

[91] Currently should at least meet the FIPS 140-2 standard

[92] Outlaw.com (2012) *ICO reiterates warning over encryption as it fines council £120k over second data protection breach,* [online], available: http://www.out-law.com/en/articles/2012/october/ico-reiterates-warning-over-encryption-as-it-fines-council-120k-over-second-data-protection-breach/ [accessed 22nd December 2014].

Business Continuity

Plans for business continuity can be more complicated and potentially less effective where data systems are distributed or responsibility is not entirely clear. The details of the data sharing scheme may need to be added into each stage of business continuity planning, from risk assessment through to practice exercises, for each organisation involved. Although potentially a small issue within the overall remit of business continuity, the receiving organisation should consider the likelihood of the supplying organisation no longer collecting the data and what impact this would have on their business.

Data validity and reliability

Understandably, ICO regards the supplying regulator as the party responsible for data validity (Information Commissioner's Office 2011: 28). This is clearly a disincentive where '…Information held by public authorities…is often partial and unverifiable' (Oswald 2013: 8). The situation is made even more difficult where the receiving regulator plans to put the data to a new use[93] (Law Commission 2014: 116). In addition, ICO expects the supplying regulator to '…establish ways for making sure inaccurate data is corrected by all organisations holding it' (Information Commissioner's Office 2011: 28). It is perhaps unsurprising that this concentration of responsibility can make it difficult for a regulator to take the decision to share its data.

Data formats and system compatibility

It is important to "…make sure that the format of the data you share is compatible with the systems used by both organisations" (Information Commissioner's Office 2011: 27). Our interview evidence suggests that this may be easier to achieve by regulators operating within a single Departmental grouping, by those fulfilling essentially equivalent functions (e.g. LA Trading Standards), or those with sufficient IT budget. Where regulators that are seeking to share use disparate systems, this is more likely to add a potentially prohibitive technical overhead.

In practice, some public authorities may continue to have difficulties sharing data even internally:

> *The LA studied was still integrating its one-stop shops, call centres, and back-office systems in order to improve the flow of electronic communications. Certainly this aspect constituted an operational obstacle hampering access to information* (Batista and Cornock 2009: 11).

Where data have to be reformatted, a data provider regulator needs to make their data approachable, reusable and supportable. In research communities, several scholars set out problems arising from reformatting for data sharing. Of those the following appear to be most relevant to the regulatory context.

(i) Inconsistency of terminology (Bache et al. 2013, Ruusalepp 2008, Hulstijn et al. 2011) with uncommon jargon, abbreviation or technical terms (Shackley 2014).[94]

[93] This could even necessitate new DPA privacy notices to data subjects or use of an alternative data source, although in practice the DPA test for this is quite wide – is the use not 'incompatible' with the original purpose?

[94] Shackley, L. (2014) 'Sharing Information: The Regulator's View', in *West Midlands Fire and Rescue Service Workshop*, Birmingham, Information Commissioner's Office,

(ii) Inconsistency of measurement units or categorical grades (Bache et al. 2013, Ruusalepp 2008).
(iii) Inclusion of unnecessary information for data re-users. In particular, confidential information that must be kept within the original data collector (Bache et al. 2013, Akers and Doty 2013, Hook International 2012).
(iv) Lost information in reformatting (Bache et al. 2013).

To address these problems, (Folinas et al. 2006)[95] suggest an advanced data formatting technology which deals with how elements in the original information are manipulated and filtered against specific rules and controls what the reformatted output looks like. Hulstijn et al. (2011) maintain the importance of common vocabularies, which allow data sharers to document their requirements, while Ruusalepp (2008) suggests coding, where possible, and templates of formatting which should be held across data sharers.

Data types

Data shared between organisations can vary between whole databases, at one extreme, and highly focused singular pieces of intelligence, at the other. According to our interviewees, the level of difficulty in sharing generally increases as one moves further towards the database end of this continuum; with greater volumes of data and more problematic issues around data subject consent.

Information shared can be quantitative or qualitative. Because many of the data sharing solutions in the research literature involve codification of data (e.g. Ruusalepp 2008), it may be that effective sharing of qualitative information - for example, responses to open questions, documents etc. - is more challenging.

Information can be general in nature (such as business name and address) or more focused on the specific requirements of a given regulator (e.g. number of recorded staff accidents). There is relatively little doubt about the efficiency savings of sharing some of the former, but the same cannot be said for the latter. According to our interview evidence and relevant literature, the value of sharing management ratings is questionable (Greenstreet Berman Ltd 2013: 12-13, 51).

As outlined above, data can be shared more easily where key factors such as data validity, formatting, metadata, comprehensibility etc. have been fully addressed by supplying and receiving organisations.

Personal or sensitive commercial information can be more difficult to share. The key factor from the perspective of the DPA is whether information constitutes ‘personal information’. If information to be shared falls outside the legal definition, it may be that there are less barriers to sharing. Even where the Act does apply, certain types of personal information may fall under DPA exemptions – e.g. s29 crime and taxation.

Understanding of data

Issues already raised in the previous sections on data formats and types highlight the importance of understanding data. Our interview evidence strongly indicates that for shared data to be useful it needs to be both clearly defined and fully understood. This is especially

[95] Folinas, I., Manikas, I. and Manos, B. (2006) 'Traceability data management for food chains', *British Food Journal,* 108(8), 622-633.

important for time-series data where data definitions can often change over time to meet changing policy requirements of the collecting organisation (see e.g. Kiddle et al. 2006).[96]

Ultimately the receiving regulator would benefit from an overall understanding of the supplying regulator. This might include their standard operating procedures, methodologies employed, aims and objectives etc. Vital to this would be an assessment of their capacity to provide valid data along with the essential metadata on sources, definitions, statistical procedures, changes over time etc.

According to a recent review, there should be far greater transparency regarding the data held by public authorities (Shakespeare 2013: 11, 17).[97] This was entirely consistent with the recommendations in the earlier *O'Hara* Review (Cabinet Office 2011: 64). Although mostly addressing the open data agenda, detailed data catalogues of, at least, core data would also undoubtedly reduce some of the technical overheads in data sharing between public authorities, a view echoed by our interviewees. However, this assumes the existence of valid and well documented datasets. Without first addressing many of the issues raised in this section of the report, it might be difficult for all public bodies to achieve this in the short term.

Responsibility, accountability and training

Calls have been made to improve training and accountability in public authorities generally (Law Commission 2013: 9). However, in common with other areas of spending, there are concerns over whether regulators can obtain the necessary resources (Law Commission 2014: 174), and where public authorities maintain shared facilities, the establishment of responsibility becomes especially important. In addition to general responsibilities for data and data systems, responsibilities and procedures for information requests[98], auditing, backups, disaster recovery etc. need to be clear.

In fact, the information assurance literature is permeated by recommendations on the organisational structure needed to safeguard sensitive data assets (e.g. Microsoft 2011: 2).[99] Relevant staff need to be fully trained and then held accountable for the data assets they own. Without the necessary training and accountability, information assurance is in danger of becoming no one's responsibility to the detriment of DPA compliance and data sharing in general (Office of the Children's Commissioner 2013: 9). This is highlighted in the latest version of the HMG Security Policy Framework (Cabinet Office 2014b: 3) and in reviews of capabilities in the science and finance sectors (Ruusalepp 2008, Akers and Doty 2013, Berman 2008).

Conclusion

The literature and interview evidence confirm that those seeking to share data can face considerable technical difficulties, and overall, it is hard to escape the conclusion that these difficulties *inter alia* may be leading some public bodies to maintain silo working. An all-encompassing 'National Data Strategy', as recommended in the Shakespeare Review may

[96] Kiddle, C., Wright, P. and Cao, B. (2006) *RSR Briefing Paper 7: The RSR Time Series Database (1989-1995) Technical Paper*, University of Cambridge: Cambridge Centre for Housing and Planning Research.
[97] Shakespeare, S. (2013) *Shakespeare Review: An Independent Review of Public Sector Information*, London: Department of Business, Innovation and Skills.
[98] Both Freedom of Information and DPA Subject Access Requests
[99] Microsoft (2011) *Business IT Security for Non-Techies,* [online], available: http://download.microsoft.com/documents/uk/smallbusiness/Business_IT_security_for_non-techies.pdf [accessed 24 September 2014].

be needed to start breaking down these barriers.[100] This could bring about a more joined-up approach to data across the public sector - encouraging strategies such as joint data sourcing and the publication of data catalogues.

The extent to which technical difficulties are regarded as insurmountable depends in part on responses to the organisational and legal aspects dealt with in previous sections. Rather tellingly, the Law Commission recently stated that '…where a [statutory] gateway is permissive, other factors may weigh against disclosure' (Law Commission, 2014: 8). Where the law leaves a choice, a decision has to be made to share or not to share. In such circumstances, technical, and organisational barriers may be cited in support of the easier option, saying 'no,' unless some of the obstacles outlined by the interviewees in this report are addressed, and incentives to share improved.

[100] Shakespeare, S. (2013) *Shakespeare Review: An Independent Review of Public Sector Information*, London: Department of Business, Innovation and Skills.

4. RECOMMENDATIONS

Organisational

- A codified duty and power for regulatory organisations to share data where this is necessary for public purposes would be almost universally welcomed by regulators. Incorporating the comply-or-explain mechanism into the code could minimise the need for legislation to address conflicts with existing legal prohibitions.

- Reinforce data sharing as the norm by including in any duty to share an obligation for regulatory organisations to publish a readily accessible register of co-regulatory, joint authority and other existing data sharing arrangements as well as future plans to improve sharing.

- It would be helpful if government departments produced a strategy and guidance on progressing data sharing from one-off short term projects through to bilateral and multilateral sharing arrangements for their reporting regulatory and other agencies, addressing in particular types of data used and the scope of those data for sharing.

- A Memorandum of Understanding template for inter-regulator data sharing would aid regulators dealing with LAs and police forces where each authority seems to insist on negotiating individually, despite having the same legal status and purpose.

- Encouraging secondments between staff involved in, or potentially involved in data sharing in both small data-poor and large data-rich regulatory organisations could be of benefit, particularly to the former.

Legal

- The powers of public bodies to engage in data sharing should be streamlined and consideration given to giving legal force to a general duty and power to share data for regulatory purposes.

- A review of UK data protection law in the area of data sharing should be undertaken along the lines recently proposed by the Law Commission, with particular focus on international best practice on modes of regulation and sanctions.

- Pending any revision of UK data protection law, the Information Commissioner should work with interested parties and sectoral bodies to develop more targeted guidance on data sharing than that currently provided through the generic Data Sharing Code.

- Public guidance on enforcement of data protection laws should address data practitioners' perceptions of a lack of proportionality in sanctions for breach of the Data Protection Act.

Technical

- An all-encompassing 'National Data Strategy' should be considered, as recommended recently in the Shakespeare Review. An initial step could be the mandatory publication of regulatory data catalogues, starting with at least core regulatory data. Further consideration should also be given to the question of whether an overarching data sharing security standard or code of practice is required.

- Regulatory organisations should evaluate the potential for greater joint or collaborative data sourcing, to gain from efficiencies and help address DPA fair processing concerns. An essential first step would be for regulators to consider the comprehensibility of their metadata and agree common definitions where feasible, prior to data collection.

- UK regulators continue to receive fines from ICO following data breaches – clear evidence that data security measures should be addressed prior to any significant expansion from current data-sharing levels. As part of any strategy to address this, regulatory organisations should publish clear lines of responsibility and accountability for regulatory data assets and improve staff training by developing a National Competency Framework for data staff.

APPENDIX 1: INTERVIEW IDENTIFIERS

Interview group **(R number = regulatory organisation identifier)**	**Data-sharing profile**	**Data-holding profile**
R6, R7, R8, R9, R10, R15, R16, R17, R31	A	ii
R28	A	iv
R1, R29, R38	B	i
R11, R12, R13, R14, R20, R21, R22, R23, R24, R25, R27, R30, R32, R35, R37, R39, R40, R41, R42	B	ii
R19	B	iii
R26	B	iv
R2, R3, R5	C	ii
R4, R34, R43	D	i
R33, R36	D	ii
R18	D	iii

Key:

A. Committed primarily to open access to data plus sharing with other regulators
B. Actively engaged in UK focused data-sharing with other regulators, some also sharing internationally
C. Actively engaged in internationally focused data-sharing but not UK
D. Not engaged actively or widely in data-sharing

i. broadly data poor/data-seeking regulator;
ii. broadly data-rich/data-supplying regulator;
iii. trade body
iv. other organisation

APPENDIX 2: INTERVIEW GUIDE

Inter-regulator data sharing (IRDS): Interview Guide Checklist

Introduction/Organisational issues

People involved in sharing data or decisions on sharing data in your organisation
Q. Who handles regulatory data in your organisation?
Q. Who is involved in decisions on the demand and supply of data?
Q. What is your interaction with others in your own organisation in respect of IRDS?
Q. Are there any champions for IRDS in your organisation, i.e. knowledgeable staff with experience in compiling, storing, and sharing regulatory information and knowledge? (Who, level, etc.?)

A little about how you see your organisation and its environment
Q. To what extent could your organisation be said to inhabit a fast changing environment and how has your organisation dealt with related change? Are data significant in this change?
Q. To what extent does informal data sharing of intelligence already exist?
Q. What other regulators inhabit your environment?
Q. Discuss own and other regulators with which you work, in terms of size, inter-organisational/inter-personal contact, and organisational risks faced etc.
Q. How would you compare your organisation's approach to data sharing with the approach taken by other regulators you have had regular dealings with?

Tell us about what you feel about IRDS
Q. What promotes/inhibits IRDS?
Q. What are the costs/benefits of data sharing?
Q. To what extent have you personally been involved in IRDS and what stimulated this?
Q. What is your own perception of IRDS in respect of costs/benefits, law, ethics, demand/supply and how does/would it affect your daily work?
Q. Do existing information policies support and encourage IRDS, or could they?
Q. Where IRDS or other DS exists, what are the factors that determine the extent to which data held are shared *(i.e. from open to closed, via, MoUs with single regulator through to DS protocols with several participants?)*
Q. Are there data held by another regulator that you don't currently have access to that could improve your compliance activities? If so, what has prevented you from accessing these data? And what benefits could these data offer to your organisation?
Q. Are there data you hold that could benefit another regulator that doesn't currently have access? Q. Do you have or can you envisage a suitable 'quid pro quo' in respect of (a) businesses and (b) other regulatory agencies – in other words do you consider your organisation is or would be a net importer or exporter of regulatory data?
Q. Are there any recent developments that have had an impact on your organisation's willingness or otherwise to share data?
Q. What one change would help increase IRDS?

Legal issues

Current and future data sharing arrangements
Q. Do you have any current Data Sharing Agreements or Memoranda of Understanding with other regulators? If so, how were legal aspects of these agreements dealt with? Describe how were any barriers to data sharing overcome? How do these arrangements benefit you?

Q. What powers are utilised to provide data to or receive data from another regulator? *(e.g. Express powers under statute; Powers implied by statutory functions; Common law powers; Schedule 15 Enterprise Act; Section of Enterprise Act (e.g. s239 – Consent, s242 – Criminal proceedings)*
Q. If you needed to set up a new data sharing arrangement with another regulator, what steps would you take?

Legal barriers to data sharing - General
Q. Do you know of any legal restrictions that could prevent you sharing data with other regulators? *(e.g. Express statutory prohibition on sharing; Lack of vires; Data Protection Act 1998; Human Rights Act 1998; Law of confidence.)*

Legal barriers to data sharing – DPA
Q. The ICO considers the DPA is quite flexible and should not stand in the way of data sharing. Do you agree? Would an improved code of practice help? If so who should be responsible for it?
Q. Is it easy to identify business cf. DPA covered personal data and if data are classed as 'personal information' under the Data Protection Act 1998, how does your approach to data sharing differ? What additional measures, if any, do you/would you want to take?
Q. Can data be anonymised in such a way as to make them non-personal and hence outside the DPA?
Q. Where data are personal, there are special restrictions for processing 'sensitive' personal data, e.g. on health, is this an issue for you?

Legal barriers to data sharing – Common law
Q. In terms of the general law on confidence, how far does disclosure to you include implied consent for disclosure to others? On what does implied consent depend?

Future legal developments
Q. What is or would be the impact of a legal requirement to share data?

Technical IT/data issues

Q. What are the easier/hardest types of data to share?
Q. What imported data are most useful?
Q. Are there any technical issues to IRDS; to what extent are data formats and definitions common across organisational boundaries (or could be made so); do you understand close other's data, and vice-versa (data holdings etc.); what are your inter-regulator networks.
Q. Do you have a specific information and data policy?
Q. Is anonymising data an issue (in respect of the DPA)?
Q. Do you have the capacity and capability to share data? If not how would you overcome this?
Q. Are you aware of datasets held by other regulators that could benefit you? Do you have sufficient information on the datasets other regulators hold?
Q. Are you/would you expect to be a net exporter or importer of data?
Q. Would greater data sharing enable long run cost savings? To whom (regulator, business, consumer)?

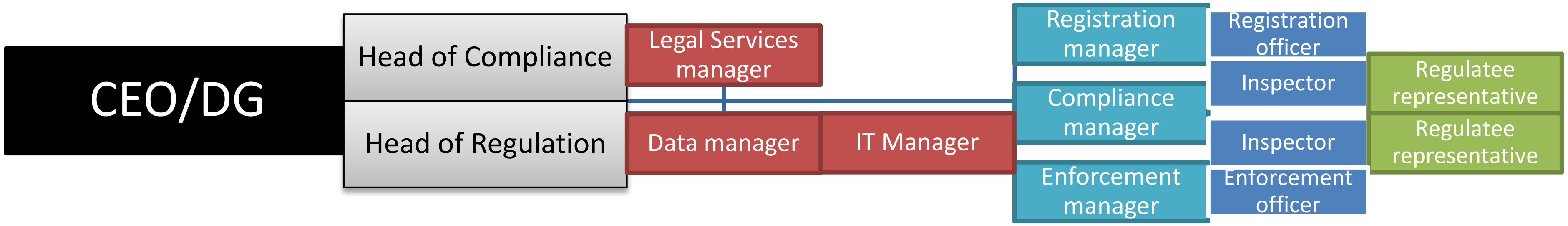
Summary

Q. Legal, technical, organisational/resource, other barriers to greater IRDS. Is it possible to rank their importance?

APPENDIX 3: INTERVIEWEE ORGANISATIONS

Organisation	Interviewee sector	No. of Interviewees
Birmingham City Council	Small retail	2
British Air Transport Association	Air travel	1
Cabinet Office		1
Cambridgeshire County Council	Small retail	1
Care Quality Commission		1
Centre of Excellence for Information Sharing		1
Charnwood Borough Council	Small retail, Food	1
Civil Aviation Authority	Air travel	2
Companies House		6
DEFRA	Food	1
Driver & Vehicle Standards Agency		10
Environment Agency	Food	3
Financial Conduct Authority		2
Food Standards Agency	Food	3
General Medical Council		1
Health & Safety Executive		5
Information Commissioner's Office		5
Intellectual Property Office		2
Legal Services Board		2
Leicestershire County Council	Small retail	1
Leicestershire Fire & Rescue Service	Small retail	1
Local Government Association		1
Ministry of Justice		2
National Federation of Meat and Food Traders	Small retail	1
OFSTED		2
Professional Standards Authority		1
Security Industry Authority	Air travel	1
The Insolvency Service		1
The Pensions Regulator		1
Westminster City Council	Small retail, Food	1

Appendix 4. Typical interviewee roles



Information Commissioner Staff

References

Akers, K. G. and Doty, J. (2013) 'Disciplinary Differences in Faculty Research Data Management Practices and Perspectives', *The International Journal of Digital Curation,* 8(2), 5-26.

Anderson, E., Tritter, J. and Wilson, R. (2006) *Healthy Democracy: The future of involvement in health and social care*, London: INVOLVE and NHS Centre for Involvement.

Ateniese, G., Pietro, R. D., Manchini, L. V. and Tsudik, G. (2008) 'Scalable and Efficient Provable Data Possession.', in *Proc. of SecureComm '08*, sprout.ics.uci.edu/pubs/a9-ateniese.pdf, 1-10.

Bache, R., Miles, S., Coker, B. and Taweel, A. (2013) 'Informative Provenance for Repurposed Data: A Case Study Using Clinical Research Data', *The International Journal of Digital Curation,* 8(2), 27-46.

Bae, J. and Gargiulo, M. (2004) 'Partner Substitutability, Alliance Network Structure, and Firm Profitability in the Telecommunications Industry', *The Academy of Management Journal,* 47(6), 843-859.

Barney, J. B. and Hesterly, W. (1999) 'Organizational Economics: Understanding the Relationship between Organizations and Economic Analysis' in Clegg, S. R. and Hardy, C., eds., *Studying Organization*, London: Sage.

Batista, L. and Cornock, M. (2009) 'Information sharing in e-government initiatives: Freedom of Information and Data Protection issues concerning local government', *Journal of Information, Law & Technology*, available: http://www2.warwick.ac.uk/fac/soc/law/elj/jilt/2009_2/bc/ [accessed 18 October 2014].

Berman, F. (2008) *Got Data? A Guide to Data Preservation in the Information Age,* [online], available: www.sdsc.edu/about/director/pubs/communications200812-DataDeluge.pdf [accessed 03 January 2015].

Better Regulation Delivery Office (2014a) *Common Approach to Competency for Regulators*, Birmingham: Better Regulation Delivery Office.

Better Regulation Delivery Office (2014b) *Regulators' Code*, Birmingham: Department for Business, Innovation and Skills.

Bock, G.-W., Zmud, R. W., Kim, Y.-G. and Lee, J.-N. (2005) 'Behavioral Intention Formation in Knowledge Sharing: Examining the Roles of Extrinsic Motivators, Social-Psychological Forces, and Organizational Climate', *MIS Quarterly,* 29(1), 87-111.

Bradford, M. (2011) *The challenges and opportunities for sharing data to combat fraud*, Newark: Regulatory Strategies Ltd.

Cabinet Office (2010) *A Strong Britain in an Age of Uncertainty: The National Security Strategy*, London: Cabinet Office.

Cabinet Office (2011) *O'Hara Review: Transparent Government, Not Transparent Citizens: A Report on Privacy and Transparency for the Cabinet Office*, London: Cabinet Office.

Cabinet Office (2012) *Categories of Public Bodies: A Guide for Departments*, London: Cabinet Office.

Cabinet Office (2014a) *Government Security Classifications*, London: Cabinet Office.

Cabinet Office (2014b) *HMG Security Policy Framework*, London: Cabinet Office.

Coase, R. H. (1937) 'The Nature of the Firm', *Economica,* 4(16), 386-405.

Constant, D., Kiesler, S. and Sproull, L. (1994) 'What's Mine Is Ours, or Is It? A Study of Attitudes about Information Sharing', *Information Systems Research,* 5(4), 400-421.

Council on Governmental Relations (2012) *Access to, Sharing and Retention of Research Data: Rights and Responsibilities,* [online], available: [accessed 22 December 2014].

Cousins, C. (2005) 'Short Term Placements: An Exercise in Organisational Culture Exchange', *Australian Journal of Public Administration,* 64(4), 81-89.

Cunningham, A. T., Bernabeo, E. C., Wolfson, D. B. and Lesser, C. S. (2011) 'Organisational strategies to cultivate professional values and behaviours', *BMJ Quality & Safety,* 20(4), 351-358.

Dasgupta, P. (1988) 'Trust as a Commodity' in Gambetta, D., ed. *Trust: Making and Breaking Cooperative Relations*, Blackwell, 49-72.

Folinas, I., Manikas, I. and Manos, B. (2006) 'Traceability data management for food chains', *British Food Journal,* 108(8), 622-633.

Frankel, M. S. (1989) 'Professional Codes: Why, How, and with What Impact?', *Journal of Business Ethics,* 8(2/3), 109-115.

Gray, J., Liu, D. T., Nieto-Santisteban, M., Szalay, A., De-Witt, D. J. and Heber, G. (2005) 'Scientic Data Management in the Coming Decade', in *ACM SIGMOD Record*,

Greenstreet Berman Ltd (2013) *Research Results: What is the Value in Regulators Sharing Information?*, Birmingham: Better Regulation Delivery Office.

Hampton, P. (2004) *Reducing administrative burdens : effective inspection and enforcement,* London: HM Treasury.

Hook International (2012) 'Information Sharing - What Can Regulators Tell Each Other', available: www.international-conference-of-legal-regulators.org/past-conferences/london-2012/information-sharing-what-can-regulators-tell-each-other/ [accessed 03 January 2015].

Hulstijn, J., Wijk, R. v., Winne, N. d., Bharosa, N., Janssen, M. and Tan, Y.-H. (2011) 'Public Process Management: a method for introducing Standard Business Reporting', in *12th Annual International Conference on Digital Government Research (DGo'2011)*, College park, MD, June 12-15, http://homepage.tudelft.nl/w98h5/Articles/ppm.pdf: http://homepage.tudelft.nl/w98h5/Articles/ppm.pdf, Independent Farming Regulation Task Force (2011) *Report: Striking a balance: reducing burdens; increasing responsibility; earning recognition*, Department for Environment, Food & Rural Affairs.

Information Commissioner's Office (2011) *Data sharing code of practice*, Wilmslow: Information Commissioner's Office.

Information Commissioner's Office (2012a) *Anonymisation: managing data protection risk code of practice*, Wilmslow: Information Commissioner's Office.

Information Commissioner's Office (2012b) 'Determining what is personal data', available: https://ico.org.uk/media/for-organisations/documents/1554/determining-what-is-personal-data.pdf [accessed 30 December 2014].

Information Commissioner's Office (2014) *Getting it right – a brief guide to data protection for small businesses, V3.1*, Wilmslow: Information Commissioner's Office.

Jarvenpaa, S. L. and Staples, D. S. (2001) 'Exploring Perceptions of Organizational Ownership of Information and Expertise', *Journal of Management Information Systems,* 18(1), 151-183.

Kiddle, C., Wright, P. and Cao, B. (2006) *RSR Briefing Paper 7: The RSR Time Series Database (1989-1995) Technical Paper*, University of Cambridge: Cambridge Centre for Housing and Planning Research.

Lacey, D. (2013) 'Predictions for 2014', *David Lacey's IT Security Blog, http://www.computerweekly.com/blogs/david_lacey/2013/12/predictions_for_2014.html,* [Accessed 22nd December 2014].

Law Commission (2013) *Data Sharing Between Public Bodies: A Consultation Paper*, Consultation Paper No 214, London: Law Commission.

Law Commission (2014) *Data sharing between public bodies: a scoping report*, London: HMSO.

Local Better Regulation Office (2011) *Data Collections: Report to the Welsh Regulators' Forum*, Birmingham: Local Better Regulation Office.

Microsoft (2011) *Business IT Security for Non-Techies,* [online], available: http://download.microsoft.com/documents/uk/smallbusiness/Business_IT_security_for_non-techies.pdf [accessed 24 September 2014].

Office of the Children's Commissioner (2013) *Office of the Children's Commissioner's response to the Law Commission consultation: Data sharing between public bodies,* [online], available: http://www.childrenscommissioner.gov.uk/content/publications/content_751 [accessed 22 December 2014].

Oswald, M. (2013) *Data Sharing between Public Bodies: Consultation Paper 214 – Response to Consultation,* [online], available: http://www.winchester.ac.uk/academicdepartments/Law/Centre%20for%20Information%20Rights/Publications/Documents/Law%20Commission%20Consultation%20No%20214%20Data%20Sharing%20between%20Public%20Bodies%20response%20from%20Marion%20Oswald.pdf [accessed 22 December 2014].

Outlaw.com (2012) *ICO reiterates warning over encryption as it fines council £120k over second data protection breach,* [online], available: http://www.out-law.com/en/articles/2012/october/ico-reiterates-warning-over-encryption-as-it-fines-council-120k-over-second-data-protection-breach/ [accessed 22nd December 2014].

Ozcan, P. and Eisenhardt, K. M. (2009) 'Origin of alliance portfolios: Entrepreneurs, network strategies, and firm performance', *Academy of Management Journal,* 52(2), 246-279.

Pardo, T. A., Nam, T. and Burke, G. B. (2011) 'E-government interoperability: Interaction of policy, management, and technology dimensions', *Social Science Computer Review*.

Pfeffer, J. and Salancik, G. R. (1978) *The external control of organizations: a resource dependence perspective,* New York: Harper & Row.

Provan, K. G., Beyer, J. M. and Kruytbosch, C. (1980) 'Environmental Linkages and Power in Resource-Dependence Relations Between Organizations', *Administrative Science Quarterly,* 25(2), 200-225.

PwC (2014) *Information Security Breaches Survey 2014,* Department of Business, Innovation and Skills [online], available: http://www.pwc.co.uk/assets/pdf/cyber-security-2014-technical-report.pdf [accessed 22 December 2014].

Reichman, O. J., Jones, M. B. and Schildhauer, M. P. (2011) 'Challenges and opportunities of open data in ecology', *Science,* 331(6018), 703-5.

Rousseau, D. M., Sitkin, S. B., Burt, R. S. and Camerer, C. (1998) 'Not so different after all: A cross-discipline view of trust', *Academy of Management Review,* 23(3), 393-404.

Ruusalepp, R. (2008) 'Infrastructure Planning and Data Curation. A Comparative Study of International Approaches to Enabling the Sharing of Research Data', available: www.jisc.ac.uk/media/documents/programmes/preservation/national_data_sharing_report_final.pdf [accessed 22 December 2014].

Sanderson, P., Seidl, D. and Roberts, J. 'Disposition to comply with flexible regulation: regulatees' perceptions of the legitimacy of codes and the comply-or-explain principle', *Regulation & Governance,* (under revision, 2015).

Schein, E. H. (1992) *Organizational culture and leadership,* San Francisco: Jossey-Bass.

Scott, D. (1999) 'Workshop on Interagency Collaboration for PANOC and DOCS staff, in C. Cousins' in [Article], Short Term Placements: An Exercise in Organisational Culture Exchange: Australian Journal of Public Administration 64: 4, 81-89.

Seidl, D., Sanderson, P. and Roberts, J. (2013) 'Applying the 'comply-or-explain' principle: discursive legitimacy tactics with regard to codes of corporate governance', *Journal of Management & Governance,* 17(3).

Shackley, L. (2014) 'Sharing Information: The Regulator's View', in *West Midlands Fire and Rescue Service Workshop*, Birmingham, Information Commissioner's Office,

Shakespeare, S. (2013) *Shakespeare Review: An Independent Review of Public Sector Information*, London: Department of Business, Innovation and Skills.

Staite, C. (2011) *Assessment of Regulatory Culture: A Literature Review undertaken for LBRO*, University of Birmingham: Institute of Local Government Studies.

Willem, A. and Buelens, M. (2009) 'Knowledge sharing in inter-unit cooperative episodes: The impact of organizational structure dimensions', *International Journal of Information Management,* 29(2), 151-160.

Williamson, O. E. (1979) 'Transaction-Cost Economics: The Governance of Contractual Relations', *Journal of Law and Economics,* 22(2), 233-261.

Wilson, J. Q. (1980) *The Politics of regulation,* Basic Books.

Yang, T.-M. and Maxwell, T. A. (2011) 'Information-sharing in public organizations: A literature review of interpersonal, intra-organizational and inter-organizational success factors', *Government Information Quarterly,* 28(2), 164-175.