Working Paper
ENGLISH ONLY

UNITED NATIONS ECONOMIC COMMISSION FOR EUROPE (UNECE)

CONFERENCE OF EUROPEAN STATISTICIANS

EUROPEAN COMMISSION

STATISTICAL OFFICE OF THE EUROPEAN UNION (EUROSTAT)

Joint UNECE/Eurostat work session on statistical data confidentiality
(Helsinki, Finland, 5 to 7 October 2015)

Topic (iv): Access to Statistical Data for Scientific Purposes

Evidence-based, context-sensitive, user-centred, risk-managed SDC planning: designing data access solutions for scientific use

Hans-Peter Hafner[1], Rainer Lenz[1,2], Felix Ritchie[3], Richard Welpton[4]

[1] Saarland State University of Applied Sciences
[2] University of Dortmund
[3] University of the West of England, Bristol
[4] UK Valuation Office Agency

Abstract: Disclosure control planning is characterised by over-reliance on theoretical models, inappropriate disclosure scenarios, worst-case planning, confusion over subjective versus objective risk management, and an unwillingness to consider the evidence base. This is most striking in the case of access to sensitive data for scientific purposes: most research on SDC has little or no value for this group. This is because confidentiality for scientific users is best managed by a range of procedural and technical options, of which statistical methods are both the least important and the least desirable.

In the last ten years or so, this procedural perspective has become increasingly dominant amongst the designers and managers of data access systems for the social sciences. However, the research management community has been less successful in getting this message out to other stakeholders.

This paper summarises the case for an evidence-based holistic approach to data access management. In particular, it considers

- the universality of the ‘intruder’ model, despite a substantial body of evidence that an ‘idiot’ model is more realistic, relevant, useful, and better aligned with legal requirements
- the focus on quantifiable measures of risk, when uncertainty is the true problem
- the legal, institutional and practical definition of ‘identification’
- assessing genuine user and stakeholder needs
- the low importance of statistical factors in the design of data access systems
- engrained institutional attitudes to risk

The common themes are use of evidence, integration of statistical and non-statistical approaches, the effective use of limited resources, and the importance of grounding strategy in realistic expectations of risk and uncertainty.

Although most relevant to the scientific research community, where the difference between worst-case and evidence-based planning is starkest, there are also general lessons for the dissemination of confidential data, for international data sharing, and even the provision of uncontrolled (public) data.

1 Introduction

Since 2013, the authors have repeatedly carried out an experiment when addressing government audiences. The audience is given a version of the following question:

> *Suppose you are responsible for making data release decisions, and have a mechanism for deciding whether such a release is 'safe'. There are two ways that you could view your role in deciding on releases:*
>
> 1. *do not release data unless the release is shown to be safe*
> 2. *release data unless the release is shown to be unsafe*
>
> *Which of these should be your default perspective?*

Functionally, the two statements are identical; but in terms of human psychology they embody two opposing views of the world. When put to audiences of data professionals, the second option (which we term 'default-open') is preferred to the first ('default-closed'), usually in a ratio of eight or nine to one.

However, when the same audience is asked a different question

> *Which of these is your organisation's default perspective* in practice*?*

the proportions are almost exactly reversed.

Default-closed is a defensive approach to decision-making, concerned with minimising costs (of a data breach) rather than maximising benefits (of useful research). It may also be a response to a perceived lack of expertise on behalf of decision-makers. A general preference for safety-first decisions in government has been widely noted (for example Buurman et al, 2012; Pfeiffer, 2008; Carlsson et al, 2012); and in the case of data access this is exacerbated by the limited benefit government departments receive from research use of data (Ritchie and Welpton, 2012).

A major influence on defensive decision-making in respect of data access is the dominance of downside risk in the literature on statistical disclosure control (SDC). SDC as a research field has existed for half a century, and the template for reporting is well-established. SDC analyses are typically carried out using 'intruder' models, worst-case scenarios, and mathematical modelling, to generate 'objective' models of release risk.

This literature is predominantly and explicitly default-closed, and so it fits well with government decision-making processes. The net result of both risk-aversion in government and defensive SDC literature is to generate a 'policing' model of data security, where right, wrong and responsibilities are clearly defined, and the aim of the data owner shifts from user needs to 'due diligence'.

The dominant SDC paradigm has much to commend it, providing a well-established template for analysing problems, particularly for comparing alternative solutions. Unfortunately, it has major flaws: with rare exceptions (eg Ronning et al, 2005), it

- does not consider non-statistical protection measures
- does not use evidence to evaluate the necessary assumptions for analysis
- confuses risk with uncertainty, leading to false perceptions of objectivity
- does not consider user needs
- does not consider the wider gains from data availability which may accrue to the data owner

Such a view of the world is likely to lead to higher costs and fewer benefits, both for society and the data owners. Moreover, this holds hidden dangers: it gives a false sense of security, and it is particularly badly suited to future developments in information technology and public opinion.

This paper argues that (a) reversing the role of user and NSI interests (b) taking a realistic approach to data security problems (c) acknowledge true uncertainty, not risk; all contribute to outcomes which are cost effective, more secure, more sustainable, more resilient, and encourage good relationships with stakeholders. We refer to this as the 'CUBEAU' model, and it reflects insights from economics, psychology, criminology, and cybersecurity.

This paper summarises the case for an evidence-based holistic approach to data access management. The common themes are use of evidence, integration of statistical and non-statistical approaches, and the effective use of limited resources. While this approach is no longer novel in some communities, it is still unfamiliar enough to cause concerns amongst those implementing data access strategies. This paper aims to address such concerns and demonstrate the importance of grounding strategy in realistic expectations of risk and uncertainty.

Although most relevant to the scientific research community, where the difference between worst-case and evidence-based planning is starkest, there are also general lessons for the dissemination of confidential data, for international data sharing, and even the provision of uncontrolled (public) data.

The paper is structured as follows. The next section summarises the characteristics of the 'traditional' approach to data access. Section three reviews the problems this causes, and section four the practical problems. Section 4 summarises the 'CUBEAU' model and gives examples of implementation. Section 5 considers the way the two models take account of future uncertainty. Section 6 concludes.

For simplicity in exposition, we shall assume that the data owners are 'national statistical institutes' (NSIs). However, the discussion is also relevant to other public and private sector data owners concerned with making data or statistical outputs available.

2 Characterisation of the dominant data access model

This section summarises some of the key features of the 'traditional' way of looking at data access as used in, for example, Hundepool et al (2012).

2.1 The conceptual basis

The traditional model of data access and SDC is default-closed. That is, it assumes that data must not be released unless it can be demonstrated to be 'safe'. Safety is determined by the application of statistical techniques to reach a pre-determined definition of ‘safety’.

Protection can be applied to the source microdata (input SDC), the statistical outputs from analysis (output SDC), or both. Input and output SDC are treated as separate problems.

2.2 The ‘intruder’

The data are being protected against an ‘intruder’: a person who has access to the data who deliberately attempts to breach protection measures, in order to glean information about one or more data subjects. This intruder is given a variety of reasons for wanting to breach data protection, for example, Eurostat (2013) lists attempting to:

- match a number of individuals with a common characteristic
- find a specific individual known to the intruder
- find an arbitrary individual to show that the published data are not secure
- take information on an individual and use this to augment external data

This intruder is assumed to have access to some external information which can be combined with the protected information. For input SDC, microdata are matched with either an external database, or the private knowledge of the intruder (for example, knowledge of a neighbour). For output SDC, intruders are assumed to have knowledge of the underlying data, and may also have access to other outputs produced from the data.

The protected data is assumed to be inherently valuable to the intruder, such that the time and effort of breaching protection brings an adequate reward.

Finally, attention is usually focused on ‘worst-case’ scenarios; this allows researchers to focus on a small number of extreme cases rather than trying to divine outcomes in many different states of the world. If a dataset is protected in the worst case (deliberate misuse by a statistical expert) it should also be protected against a much weaker case (a member of the public finding an unencrypted CD left behind on a train).

2.3 Technical support

The broad principles of SDC are relatively stable; although advances are made, at any point in time there is a well-established perspective on the pros and cons of various methods, and so data owners can employ methods which have been extensively described and analysed. In 2009 Eurostat commissioned a major project to review, summarise and recommend SDC best practices for NSIs. The resulting project report and book (Hundepool et al, 2010b, 2012) provides a comprehensive overview but does not differ substantially from similar works published over the last fifteen years or so.

This has encouraged the development of software libraries, as well as complete general-purpose packages. The Eurostat-sponsored programs τ-Argus and μ-Argus, providing tabular and microdata protection respectively, were developed with the European Statistical System in mind, and are recommended to all European NSIs. The R packages sdcTable and sdcMicro are also increasing in popularity. A decision-maker can, to a large degree, pull ready-made SDC solutions off the shelf.

2.4 The strengths of the traditional perspective

The traditional model has a number of advantages:

- *Familiarity*: the widespread agreement on how disclosure scenarios are conceptualised, specified and analysed means that research shares a common frame of reference; discussions can focus on new developments and critiques of existing models

- *Reproducibility*: a standard way of defining disclosure scenarios simplifies the replicability and comparability of critical analyses
- *Accumulated expertise*: the direct application of many theoretical tools to practical outcomes means that theory can be informed by experience
- *The level playing field*: a common approach to scenario specification reduces the opportunity for researchers to design tests to favour one solution over another; as a result, there is often widespread agreement on which SDC procedure work best in particular situations
- *Envelopment of alternative models*: the focus on worst-case scenarios reduces the set of scenarios to be considered

It is clear that there is immense value in the body of knowledge built up on SDC techniques. The purpose of this paper is not to argue that SDC theory is of no value; instead we argue this value is lost because of the indiscriminate use of such tools in inappropriate environments.

3 Problems with the traditional model

3.1 Default-closed versus default-open

The traditional perspective has a default-closed perspective, as all the literature uses the language of "release *if*…" From a theoretical perspective, there is no reason for choosing the default-closed over default-open; the default perspective should not change the achieved outcome in a specific practical situation. There is an analogy with the dual/primal characterisation of linear programming: an economic model specified as "maximise x subject to y" that is reparameterised as "minimise y subject to x" returns the same optimal solution.

However, the analogy is not exact. In linear programming, values are known with certainty; in data access decisions, most of the true 'parameters' are unknown and hence evaluated subjectively. Moreover, humans values losses and gains differently. The default-open model starts from "full utility; no security", whereas default-closed starts from "full security; no utility"; each model therefore will see different gains and losses when moving away from the default position[1].

Ritchie (2014) uses the diagram in Figure 1 to summarise the importance of the default:

[1] 'Full' and 'no' are illustrative simplifications: collection methods mean unrestricted survey data still have some minimal level of protection; unreleased data can still provide value by bespoke services.

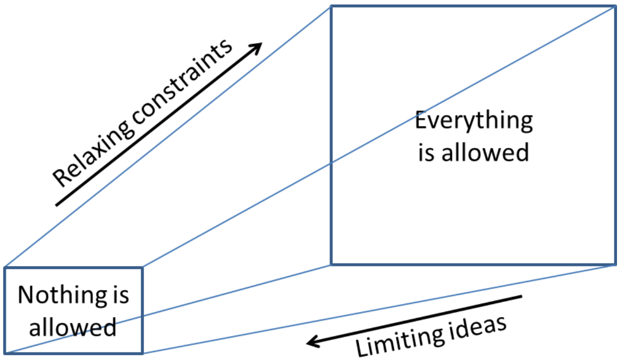

Figure 1 The importance of the starting point (from Ritchie, 2014)

Moving from default-closed, on the left, emphasises the loss of constraints. Moving from default-open, on the right, emphasises the loss in utility. Seen from this perspective, it is clear that the two data owners with different starting positions would be unlikely to agree on an appropriate balance of security and utility; ceteris paribus, default-closed leads to less data access.

There is no theoretical reason for preferring default-open over default-closed or vice-versa, although it could be argued that non-data professionals concerned by the potential data privacy violations would always prefer default-closed. However, even amongst data professionals, the preference for default-open noted in the Introduction is largely an abstract concept; in practice data professionals apply default-closed. This is also invariably the position adopted in research papers because it simplifies comparability of results, but this means a convenient convention is unnecessarily restricting data access.

3.2 Private risks, public benefit and the shifting of responsibility

There is a second reason why governments prefer default-closed. As several authors have noted (eg Lofstedt, 2004; Yang and Holzer, 2011; Ritchie 2014), decision-making in public life is often “specific cost, diffuse benefits”: many will benefit from a good decision, but one person will be held responsible for a bad decision. Ritchie and Welpton (2012) note that research data access is a good example of this: an NSI typically expects little direct benefit from the research use of its data, but expects it will face most of the criticism if a breach of confidentiality occurs. Hence, it is rational for an NSI to seek to restrict data access as far as possible. This is unlikely to be good for society as whole.

When an NSI allows access to data, risk can be limited by shifting responsibility for system failure onto the user. Consider a shopkeeper who places obvious signs in his shop that “thieves will be prosecuted”; is he to blame if a thief chooses to ignore the obvious signs? Clearly not. Thus also for NSIs: if due diligence is undertaken to show that the correct operations of systems, and the correct level of data protection, is in place, then it follows that any breach of confidentiality is the fault of the user. The role of the NSI is to ensure that the user is fully aware of his or her responsibilities; the NSI cannot be held responsible if the user wilfully ignores them.

This model only makes sense in a world of certainty populated by (boundedly) rational individuals. The shoplifting analogy, while appealing to NSIs, is not appropriate to data access situations. A better analogy is with software licences. Most humans click on “accept terms and install” without bothering to read the T&Cs; the software companies believe this gives them legal protection. Similarly, an NSI may require a user to read guidance on safe behaviour; if the user doesn't read it and commits an offence, it's the researcher's problem. This leads to 'them and us' perceptions amongst NSIs.

Unfortunately for NSIs, courts have increasingly struck down software companies' ability to rely on unread T&Cs; it is unlikely that privacy commissioners would be more lenient to NSIs that similarly assume that distribution of information leads to acquisition of information by users.

In summary, NSIs are incentivised to limit data access more than society as a whole would prefer. NSI create a 'them and us' position by trying to shift responsibility for any data breach onto users, even though it is quite likely that faith in the transfer of responsibility is unjustified.

3.3 The lack of intruders

As noted above, the traditional SDC literature is entirely focused on the concept of the 'intruder'; and there are advantages to comparative analysis from the theoretical model of the sophisticated statistician willing to expend unlimited time acquiring corroborative information and then cracking SDC protection. This has some validity when considering unrestricted access to the data; for example, there are cases of individuals from pressure groups attacking tables published by NSIs; and if a public use microdata file is to be produced, it is sensible to suppose that an individual may try to breach confidentiality just to be mischievous.

However, this does not fit the scenario of highly motivated experts with unlimited time and enthusiasm. Moreover, in the case of data access limited to bona fide researchers, there is **no** evidence to support the 'intruder' model. There are examples of researchers doing foolish, and sometimes unlawful, things; but no verified[2] examples of a researcher re-identifying data for gain in the way specified in intruder scenarios.

3.4 Infeasible scenarios

Practical studies refute most of the assumptions about the ease of breaching protection. Repeated studies (eg Lenz, 2006; Hafner, 2008; Evans and Ritchie, 2009) have shown that external databases provide very poor matching opportunities for microdata, even where direct identifiers exist (for example, in company data). There is little or no evidence that disclosure by dominance exists in practice, requiring as it does very high knowledge of contributor data.

Flaws in SDC protection have been identified by lobby groups and 'black hat' analysts, but never leading to the unacceptable outcomes specified in disclosure scenarios. In perhaps the best-known example, UK abortion statistics were interrogated to reveal the identity of one doctor (amongst several) who performed late-stage abortions; however, none if the patients was identified. Moreover the case came

[2] 'no verified': the authors are aware of two anecdotal cases of deliberate re-identification by researchers. In one, it is not clear whether the re-identification plan was carried out; in the other, it appears that data with direct identifiers was analysed. The circumstances in both cases are uncertain, as the sources had only second-hand reports of the incidents, and were not able to verify the reports. To put this in context, the authors have, between them, worked with NSIs and data archives on every continent, and have been actively seeking information on breaches of confidentiality. Where verified, these have always been the result of breach of procedure, and never attempts at re-identification.

to court as the lobby group responsible was unable to identify any personal information from the published data (McHale and Jones, 2012)[3].

Finally, it is not at all clear that the value of the uncovered data warrants the effort required and the legal risks, at least for NSI data. For example, Buccellato and Ritchie (2010) reviewed the commercial value of ONS' business microdata, and argued that this was negligible as much more timely, accurate, relevant information was available from external sources.

3.5 Mathematization of risk measures

Despite the infeasibility of the scenarios, these are used to generate risk measures. All microdata assessments describe the characteristics of the data, and then postulate a potential 'matching database' with particular characteristics.

One difficulty is that the characteristics of such a database are not known; hence, the matching database usually reflects the most problematic elements of the source database: population uniques, extreme values, sparse sets. This is to ensure that worst-case scenarios can be accounted for. An alternative, used for example by μ-Argus software, is to treat the source database as the matching database, in the absence of any other information. While this is practically convenient, it makes no sense as a measure of risk. Moreover, there is strong empirical evidence (eg Lenz, 2006) that these worst-case scenarios strongly over-estimate the real risk of identification.

A further difficulty is that risk models are variable dependent: measures are based on a vector of key variables, but when these change, risks change. As well as relying upon agreement on the key variables, an NSI trying to satisfy multiple requests could end up by creating multiple versions of the same file which independently are 'acceptable' to release, but when put together are problematic.

Similarly, for tabular outputs there are models (and software) to analyse the risk of such outputs before and after table protection; as a measure of the impact of SDC they can be useful, but claims that tables are now 'safe' cannot be upheld as it is not possible to prove that disclosure by differencing cannot happen.

Such mathematical models do a useful job of evaluating the *relative* risk reduction of alternative disclosure methods, but they have little validity as measures of *absolute* risk. Theoretical articles, descriptions of risk measures, and user guides for software tools do not usually claim absolute validity. Nevertheless, human psychology tends to place more value on mathematical measures, and so such measures will influence decisions irrespective of their validity. Again, the problem is not with the tools themselves, but the way they are employed to generate spurious precision.

3.6 Risk versus uncertainty

The focus on mathematical modelling obscures a more fundamental truth: that, although the SDC literature discusses the 'risk' of disclosure, this is not a true risk at all. It is Knightian uncertainty. Knight (1921) distinguished between *risk* as a measurable mathematical value ("what is my probability of getting cancer?) and

[3] The court decided that the doctor did not need the same level of confidentiality protection as the patient as public interest required knowledge of the number of abortions requiring special exemptions, and the doctors expected to be publicly accountable for their actions. This did not mean that the doctors should be identified, only that they did not require the same level of protection as patients. See McHale and Jones (2012), and references to court proceedings therein.

fundamentally unmeasurable *uncertainty* ("will new cancers be discovered?"). In the context of SDC it is clear that uncertainty is the relevant concept.

Consider the conditions necessary for a deliberate breach to occur:

- An intruder has sufficiently strong motives to breach SDC protection
- The intruder has the necessary time, patience and skills to do so
- The intruder has sufficient additional information to help him or her do so

We have no evidence whatsoever to allow us to judge the 'likelihood' of any of these conditions. These are not 'black swan events' – unforeseeable events which are rationalised after the fact (Taleb, 2010). The scenarios are perfectly foreseeable – we just have no idea how to allocate a probability to them.

3.7 Subjectivity

SDC is presented as an objective process which can provide the correct answer, albeit with some subjective decision-making about uncertainty at the limits. This perspective supports the NSI's decision to (1) rely upon SDC measures to protect the data, and (2) shift responsibility for any remaining risk onto the users.

However, it should be clear that the true situation is the exact opposite. Knightian uncertainty and the substitution of practical risk models for realistic scenarios mean that SDC decisions are almost entirely subjective. As Skinner (2012) notes, even the 'objective' mathematical models are largely determined by subjective choices about parameters, scenarios, and ease of analysis.

Acknowledging the subjectivity of decision-making implies recognising the potential validity of alternative viewpoints. This is incompatible with the traditional model, with its emphasis on black and white definitions of safety and responsibility.

3.8 Inappropriate focus

The focus on intruder scenarios creates a second problem; it distracts attention from known problems in data access. Data professionals are aware of breaches of procedure, sometimes leading to breaches of confidentiality. These are overwhelmingly due to two causes:

- mistakes
- users finding ways to avoid tedious, time-consuming or otherwise unappealing procedures

These are known problems, and yet almost no analysis has been spent on this issues (with the exception of the RDC community; see below). Instead data access management generally

- does not ex ante acknowledge the likelihood of error
- considers user misbehaviour to be a breach of the user's responsibility to act in accordance with the expectations of the NSI[4]

This is entirely consistent with the black-and-white, responsibility-shifting traditional model; but it effectively means that the major source of confidentiality risk is not being addressed.

[4] One author recently pointed out to a data control authority that its procedures were being actively bypassed, and received the response "Well, that's wrong; they must follow the rules. That's [the organisation's] policy, and everyone should know it."

3.9 Legal, institutional and other definitions and guides

Terms such as ‘confidential’, ‘identified’, ‘disclosive’, ‘personal data’ and so on are difficult to define theoretically, let alone in practical situations. ‘Non-disclosive’ is particularly difficult, as it is non-provable; it is only falsifiable.

Most laws reflect this ambiguity. Decisions on what counts as ‘sensitive personal data’, for example, are usually delegated to some authority to consider in respect of specific cases. Additionally, most laws include some measure of ‘likelihood’ or ‘reasonable risk’ to allow for the impossibility of guaranteeing non-disclosure; for example, German law specifically references ‘de facto’ anonymisation.

In the authors’ experience, all data professionals are aware of these ambiguities in definition and law; and yet these are rarely referenced in data access discussions. Instead the debate is framed in terms of what is allowed rather than what is likely to be acceptable.

Ignoring ambiguity and the nuances of language or statistics can be essential for effective decision-making; without arbitrary decisions on what is or is not ‘disclosive’, agreement on solutions may be impossible. The concern is that, over time, an accepted convention can become a ‘fact’. The traditional model, with its emphasis on the existence of a ‘right’ solution, discourages recognition of this ambiguity.

3.10 Role of the user

The user is largely absent from the SDC literature. Lip service is often paid to maintaining the ‘utility’ of the data, but it is clear that utility is a measurable outcome of the data protection process, not an objective. Whilst there are exceptions such as ISTAT (2013), the overall impression from the SDC literature is that users should be grateful for any crumbs of data thrown to them. For example, almost all measures of the ‘cost’ of data protection focus on the univariate characteristics of the perturbed data; these are easy to measure and compare. Unfortunately, users are likely to be interested in marginal analysis, and multivariate relationships are much harder to maintain.

As noted above, most data breaches occur through the actions of disengaged users who have no interest in the NSI’s welfare. The likelihood of this is increased by failing to put users at the centre of decision-making.

Ignoring the user also means that opportunities to make data safer might be unexploited. For example, if the only purpose of the data is to carry out linear regressions, than a cross-product matrix provides all the necessary moments for analysis; there is no need to release individual microdata (see Ritchie 1995).

Finally, treating the user as a problem rather than a solution means that the NSI is missing out on the gains that come from a positive relationship with researchers: free methodological input, expert user feedback, leveraging of investments in data collection, and so on. It was noted earlier that NSIs often see data access as a cost with little benefit; ignoring the user almost guarantees that the benefit is minimised.

3.11 Non-statistical factors in system design (either/or approach)

The SDC literature answers a very specific question: “if the data or outputs were to be released to users in this way, what would be necessary protection to prevent breaches of confidentiality?” In practice, the question becomes “If we apply these protection measures, can this data be released to this group?”

The difference is subtle, but important. The first question emphasises the residual nature of SDC protection: what must be done to protect the data given all the non-statistical protections that could also be employed. Given that all data protection reduces the value of the data, this seems to generate the ideal solution for users.

In contrast, the second phrasing emphasises the primacy of the SDC protection. The SDC process has given an answer to the question of "how safe are these data/outputs?", and all that remains to build a distribution mechanism using non-statistical measures. This fits neatly into the traditional model's preference for unambiguous decisions about safety.

3.12 Problems with the traditional approach to SDC: a summary

The SDC literature is wide and deep; for fifty years it has provided a coherent, consistent canon of theory providing insights into various statistical approaches to data protection, both of inputs and outputs. However, when applied to practical problems of data access there are four substantial flaws:

- the scenarios used in theoretical discussions have little basis on reality, limiting their applicability
- the use of mathematical language to express and analyse problems provides a false sense of objectivity and measurement
- almost all discussion assume a measurable risk , whereas Knightian uncertainty is the relevant concept
- the data centred approach leads data managers to overestimate the risks and underestimate the benefits from non-statistical approaches

4 Developing a new approach: 'CUBEAU'

In recent years, an alternative approach to data access has developed, which for simplicity we will refer to as 'CUBEAU' (Centred on Users; Based on Evidence; Acknowledging Uncertainty). This approach is increasingly accepted as best practice for research data centres (RDCs), and recently has started to have a wider impact on planning. In this section we describe the conceptual basis, and then give examples in practice.

4.1 Principles of a revised approach

The CUBEAU model is empiricist in outlook, but more interpretivist than traditional statistical modelling. The main elements are as follows

- Identification of objectives derives from the analysis of user needs; data owners' confidentiality concerns are constraints on the unlimited distribution of data, not objectives
- Attitudes to data release are default-open
- The access environment is fundamentally uncertain in the Knightian sense; use of evidence can reduce, but not eliminate, such uncertainty
- Risk assessment uses an 'idiot' model of researcher behaviour, characterised by lack of knowledge and poor decision-making; 'intruders' are an extreme subset
- Primary risk assessment is based on (1) empirical evidence of known breaches (2) psychological, institutional and commercial risk factors
- Risk scenarios are subjectively weighted by empirical plausibility and impact

- Statistical and non-statistical approaches both contribute to provide safe access, but the relative contribution of each depends upon the specific type of data release
- Where feasible, non-statistical solutions are preferred over restrictions on data or output (ie SDC is a residual)
- Modelling tools provide supporting information on relative risk assessment, but not absolute or definitive guides
- Decision-making is acknowledged to be subjective; the data owner's right to make decisions should not be confused with the decision to be right
- Major errors are possible; minor errors are certain to occur; planning should acknowledge this
- Changing the environment is part of the decision-making process: in particular, attitudes of all parties (especially researchers) can be changed through training and education

4.2 Advantages of the CUBEAU model

The CUBEAU model tends to produce access solutions with lower costs, better security and improved user acceptance.

Costs are lower than in data-centred solutions because CUBEAU takes account of statistical and non-statistical factors and treats the environment as a factor to be altered. Costs are also lowered by using evidence to focus on plausible outcomes, rather than extreme cases with no realistic chance of happening. For example, the first fully-CUBEAU-compliant RDC, the UK Office for National Statistics' Virtual Microdata Laboratory (2004-2011), had operating costs 'an order of magnitude' lower than comparable systems across Europe, yet still supported more researchers with more data and faster response times.

Security is also improved by combining statistical and non-statistical factors: known non-risk factors (such as researcher error) are tackled within an integrated framework. It seems likely[5] that the focus on uncertainty, the expectation that errors will be made, and the acknowledgement of the subjectivity of decision-making means that CUBEAU models are better prepared to manage risk as an ongoing problem (the data-centred model sees risk as something to be addressed by SDC and/or system approval). For example, in one RDC a researcher was penalised following a breach of procedures. No changes were made to the RDC operation: it had been acknowledged when the RDC was set up that such a risk could occur, but making it impossible would have made the system unworkable; the manifestation of the risk had not changed that assessment. Ironically, this breach of procedure strengthened the CUBEAU approach for the facility: the negligible impact of the breach was entirely in line with predictions.

Finally, the CUBEAU model focuses on the user experience as its objective, and it also accepts that the data owner has no monopoly on truth; hence, it tends to produce data access solutions which are more aligned to user needs. Disgruntled users lead to both higher costs (as they do not contribute to the system security) and higher risks (researcher attitude is a known risk factor), and so the user focus contributes to lower cost and better security. Giving researchers a stake in the system can make them self-policing. The basis for this is well established in psychology and criminology (see Ritchie and Welpton, 2015), and the evidence from data access supports this view. Examples of this in practice include researchers

[5] It is, of course, impossible to be certain…

- reporting their own errors
- reporting the errors of others
- developing additional self-training within peer groups

4.3 Practical implementations

The main practitioners of the CUBEAU approach have been research data centres (RDCs) holding government socio-economic data. This is for a number of reasons:

- RDCS by design do very little to restrict the data, and so most of their protection comes from non-statistical factors
- the 'intruder' model has no relevance to the RDC case but there is widespread understanding of user errors
- traditional approaches to output SDC (almost exclusively devoted to tabular data protection) do not address the questions raised by research outputs
- there is a clear cost gain from self-policing researchers, particularly for output checking
- most RDCs carry out some researcher training anyway, and so amending this to reflect the CUBEAU ethos is relatively straightforward
- the original CUBEAU model developed for UK RDCs was extensively documented, allowing similar facilities to adapt to their local purposes and to claim precedent (which is important for persuading risk-averse data owners)

CUBEAU-style models have been adopted by RDCs across Europe, north and central America, and Oceania[6]. Not all centres adopt the full approach, but the common themes are strong user-centred focus, awareness of the importance of user training, and the use of SDC to supplement non-statistical tools.

When considering making data safe for distribution, CUBEAU has made less impact to date because of the weight of SDC literature supporting the traditional model. The software package mu-Argus has a weighty user manual providing advice on the tool's use and the statistical theory behind it, but almost no guidance as to how to interpret the results from a non-mathematical standpoint. The R package sdcMicro provides a similar functionality; however, new use guidelines being developed by the UK Data Service are designed to place the use of SDC micro in a CUBEAU framework, with the role of subjective set-up and interpretation more clearly defined.

Despite the lack of application so far, the gains to anonymised distribution can be substantial. Hafner et al (2014) applied the CUBEAU approach to creating a Scientific Use File from the 2010 Community Innovation Survey, business microdata from seventeen European countries. The traditional approach, used on previous waves, independently perturbed 100% of observations for all continuous variables, yet did not address identification through dominance. In contrast the CUBEAU implementation perturbed under 1% of the observations for just one continuous variable and explicitly maintained relationships between variables – important as marginal analysis is the main value of the microdata. It should be noted that the previous method following best practice according to traditional convention – but by failing to address user needs and non-statistical factors it produced a dataset which was viewed by users as useful for teaching only.

[6] The 'five safes' planning framework is widely referenced, but this a design tool within the CUBEAU framework rather than a philosophical position. In the US, the 'portfolio' model of eg Lane et al 2008 is widely referenced; it is a near-synonym for the Five Safes, and shares many CUBEAU characteristics.

Outside of specific access solutions, the CUBEAU framework has begun to influence strategic thinking, albeit on a country-by country basis. This is most evident in the UK, where planning (if not always implementation) is increasingly framed in the CUBEAU language. In recent years the agriculture, tax and justice ministries have all produced and implemented user-centred, evidence-based research data strategies. In addition, the academic funding bodies have challenged the Information Commissioner's data-centred perspective in favour of CUBEAU-style risk assessment. The CUBEAU approach is also spreading beyond the socio-economic community into health research (eg Wellcome, 2015), with epidemiological research centres starting to adopt the operating models of the social scientists.

Outside the UK, the traditional model still dominates, with individual exceptions: for example, ISTAT has been an outlier for some years in taking a genuinely user-centred approach; in Germany, researchers are exploring linked input-outputs SDC in a modern framework.. While the NSIs in some countries, such as New Zealand and Australia, have formal strategies recognising the CUBEAU concerns, in other countries the impact tends to be determined by the influence wielded by the major RDCs. An interesting development is the recent willingness of Eurostat to explore and experiment with CUBEAU principles; given Eurostat's position in the European Statistical System, this has the potential for a significant change in culture across a range of countries.

5 Future-proofing data access

The rise of personal information placed on social media, 'big data' collected by public and private organisations, increasing use of administrative data, unpredictable new sources of data such as Twitter, Google Glass or internet-enabled fridges… All of this means that the future data collections by government statisticians are likely to be different from current ones; beyond that, there are no real certainties.

The traditional SDC model was developed while NSI data was almost exclusively survey data or collected from a limited number of administrative sources. Sixty years of unchanged data collection methods allowed NSIs to develop elaborate scenarios based upon a limited range of actors and external data sources. This relatively simple and stable world well suited the policing approach of the traditional model, where rules were set in advance and rarely reviewed.

Future developments provide two major challenges to this perspective. First, with an increasing range of data sources, the likelihood increases that the 'confidential' data held by the NSI is known and accessible to others. For example, if all data is collected from administrative sources, then by definition all of it is available to a third party; moreover, the third party is likely to have fully identified data, not just the de-identified or anonymised data supplied to researchers.

Second, being able to predict what is a likely matching scenario (let alone a worst-case one) becomes an exercise in speculation. The use of 'black hat' analysts to find flaws in NSI data protection strategies have shown that even in the traditional areas of NSI tabular output, previously acceptable strategies may no longer be sufficient [ref Gelbert stuff].

Combined, these two fundamentally change the intruder scenarios. A scenario where an intruder knowingly has the exact same data as the NSI but with full identifiers, directly linkable to a near-infinite range of other data sources, is much more extreme than current models. Meanwhile, RDC managers should be taking measures to counteract wearable cameras and ensuring that biometric tokens can't be forged.

The CUBEAU approach does not, of course, solve such problems; but it does put them in context. For example, we know that individuals with access to administrative data sources do use them to breach confidentiality, but they do not do this by reverse engineering NSI products; why would they? Similarly, there is no rule or technical precaution that can prevent an RDC user simply looking at an interesting observation and remembering it[7]. An RDC manager believed that there was a significant likelihood of researchers secretly filming, for example, should be questioning why such researchers are allowed access to the data at all; more policing is unlikely to be the answer.

Focusing on data protection is a risky strategy when it appears that the possibilities for breaching confidentiality can only increase. A more productive approach would be to focus on the likelihoods of such breaches, and the incentives for doing so – and this is the core of the CUBEAU approach.

The CUBEAU emphasis on uncertainty also helps future planning. CUBEAU planning explicitly highlights the likelihood of failures of any system, and acknowledges that perceptions of 'best' practice change over time. This is an attitude that says: failures can't be prevented, but accepting that everything we do may turn out to be wrong is a good way to prevent most of it being wrong.

6 Conclusion

The traditional model of data access, focused on data, policing, and formal tools for SDC, has many strengths. Amongst these, the most important is a good understanding of how different SDC techniques fare in the face of different confidentiality problems. This is achieved by having consistency of treatment: similar attack scenarios and external data sources, similar univariate measures of utility, similar assumptions that a breach per se is unacceptable, and so on. This consistency is essential for developing comparability amongst the broad corpus of models.

The concern is that these models have been applied in practical situations of data access with little recognition that they are theoretical constructs. They have serious flaws when taken as a guide to action: there is no evidence supporting the 'intruder' model for academic researchers, and very little even for public data outputs; models of 'utility' are largely irrelevant for researchers; no estimation is made of the value of the data, or of the genuine expect impact of confidentiality breaches; little or no use is made of the evidence provided by half a century of research data access.

However, the two biggest problems are the context of the SDC itself. First, SDC takes place in fixed environment: the only choices to be made are about the data protection, and we assume, as nothing else can be controlled, that worst-case scenarios will occur. Second, access is default-closed; protection of the data is the desired outcome, utility the constraint.

[7] Remote-job system managers have argued this what their system avoids, but it is straightforward to show that it is not possible to build a system which is both risk-free and useful; experience also shows that researchers are very good at identifying flaws in such systems. As for RDCs, the real question is: if you think your researchers are going to put in so much effort to break your system, why are you letting them have access?

This model appeals to government decision-makers, who are naturally risk-averse, and who also face incentives encouraging a safety-first attitude. The traditional model leads to a policing approach, with the NSI carrying out due diligence and everything else being researcher's fault. Rules are set and are there to be obeyed.

There is more than enough evidence now to argue that this approach is fundamentally counter-productive: if provides a false sense of security, at a higher cost and lower benefits, and makes enemies of the NSI's natural partners.

This paper has summarised an alternative perspective, characterised as centred on users, based on evidence and acknowledging uncertainty (CUBEAU). This approach differs from the traditional perspective at all levels. Conceptually, it reverses the traditional default-closed model, treating user needs as the objective and data protection as a constraint. Epistomologically, it takes a more interpretivist approach rather than the pure positivism of SDC, acknowledging the subjectivity of decisions and the evolution of knowledge. Practically, it insists upon the use of evidence to evaluate risk scenarios; basing decisions on likelihood rather than possibility; considering non-statistical measures along with statistical measures; and managing risks in the expectation that some at least of those risks will manifest themselves. Underlying this is the recognition that this decisions in this areas are taken in an atmosphere of genuine uncertainty.

This world view is much less comfortable than the traditional view. It is uncertain, subjective, empirical, and affords no opportunity to shift responsibility onto others. But these are also its strengths: it is flexible, willing to change, pragmatic, and seeks to bring parties together to share common goals, where possible.

This paper does not advocate abandoning the great store of knowledge built up on SDC, but it argues that these SDC models can be made much more flexible and adaptable when seen as just a part of a much wider range of tools. Putting them in a better context, making balanced subjective judgement based on evidence, and putting users first leads to a win-win outcome: lower costs, better security, better data and outputs. While this has already been demonstrated for restricted-access research data, the broad principles seem likely to also hold true for unrestricted data files and outputs.

Perhaps most importantly for decisions makers, recognising the uncertainty in decision-making is a key step in addressing potential risks. The traditional model reassures decision-makers that risks have been managed; the CUBEAU approach reminds decision-makers that, at best, a fair balance of uncertainty has been achieved. The latter is, we believe, a more intellectually honest approach, and provides a better framework for managing that uncertainty.

Finally, the traditional model is under severe pressure from developments in data availability, providing problems to which SDC models have no obvious answers. By emphasising uncertainty, change, user value, and the use of full range of tools available to manage data access, the CUBEAU approach already has the framework in place to address any questions the world can throw at it.

References

Buccellato T. and Ritchie F. (2010) *Commercial value of ONS business microdata*. Report for Statistical Policy Committee. Office for National Statistics

Buurman M., Delfgaauw J., Dur R. and van den Bossche S. (2012) *Public sector employees: Risk averse and altruistic?*, CESifo Working Paper: Behavioural Economics, No. 3851 http://hdl.handle.net/10419/61046

Carlsson F., Daruvala D. and Jaldell H. (2012) “Do administrators have the same priorities for rtisk reduction as the general public?”. *J. Risk and Uncertainty* v48 pp79-95

Eurostat (2013) *Guidelines for anonymisation of social survey microdata*. Luxembourg. Eurostat

Evans, P. and Ritchie, F. (2009) *UK Company Statistics Reconciliation Project: final report*, Report for the Department of Business Enterprise and Regulatory Reform; URN 09/599

Hafner, H.-P. (2008). *Die Qualität der Angriffsdatenbank für die Matchingexperimente mit den Daten des KSE-Panels 1999 – 2002*. Mimeo. IAB

Hundepool, A., Domingo-Ferrer, J., Franconi, L., Giessing, S., Lenz, R., Longhurst, J., Schulte Nord-holt, E., Seri, G. and De Wolf, P. (2010). *Handbook on Statistical Disclosure Control*, ESSNet SDC http://neon.vb.cbs.nl/casc/.\SDC_Handbook.pdf

Hundepool, A., Domingo-Ferrer, J., Franconi, L., Giessing, S., Schulte Nord-holt, E., Spicer K., and De Wolf, P-P (2012) *Statistical Disclosure Control*. Wiley Series in Survey Methodology.

Knight, F. H. (1921) *Risk, Uncertainty, and Profit*. Boston, MA: Hart, Schaffner & Marx

Lane, J., Heus, P., & Mulcahy, T. (2008). Data access in a cyber world: Making use of cyberinfrastructure. *Transactions on Data Privacy*, 1, 2-16.

Lenz R. 2006. Measuring the disclosure protection of micro aggregated business microdata - an analysis taking as an example the German Structure of Costs Survey, *J. Off Stats* 22 (4):681-710

Lofstedt R.E. (2004) “The swing of the regulatory pendulum i Europe: from precautionary principle to (regulatory) impact analysis”. *J. Risk and Uncertainty* v28:3 pp237-260

McHale J. and Jones J. (2010) "Privacy, confidentiality and abortion". *J Med Ethics* 2012 38: 31-34 doi: 10.1136/jme.2010.041186

Pfeifer C. (2008) *Risk aversion and sorting into public sector employment*. IZA Discussion Papers no 3503. http://ftp.iza.org/dp4401.pdf

Ritchie F. 2014a. Access to sensitive data: satisfying objectives rather than constraints, *J. Off Stats* 30(3):533-545, September. DOI: 10.2478/jos-2014-0033

Ritchie F. 2014b. *Resistance to change in government: risk, inertia and incentives*. Working papers in Economics no. 1412, University of the West of England, Bristol. December

Ritchie F. and Welpton R. (2012) “Data access as a public good” in *Work session on statistical data confidentiality 2011*, UNECE/Eurostat.

Ronning, G., Sturm, R., Hoehne, J., Lenz, R., Rosemann, M., Scheffler, M., and Vorgrimler, D. 2005. *Handbuch zur Anonymisierung wirtschaftsstatistischer Mikrodaten*. Statistik und Wissenschaft, Band 4, DeStatis

Taleb N. (2010) *The Black Swan: the impact of the highly improbable* (2nd ed.), London: Penguin,

Yang K. and Holzer M. (2006) “The Performance–Trust Link: Implications for Performance Measurement” *Public Administration Review*. v66:1 pp114–126 http://onlinelibrary.wiley.com/doi/10.1111/j.1540-6210.2006.00560.x/pdf