

Etiquettes of Data Sharing in Healthcare and Healthcare Research

M. Hartswood¹, K. Ho¹, R. Procter², R. Slack¹, A. Voss¹

¹Social Informatics Cluster, School of Informatics, University of Edinburgh

²National Centre for e-Social Science, University of Manchester

Corresponding email address: mjh@inf.ed.ac.uk

Abstract. Our studies suggest that data is not an unencumbered commodity but that it comes entangled with obligations: sharing data reflects a nexus of rights, obligations and expectations associated with collegiality, data ownership, confidentiality, IPR, competitive advantage, ethics, and other organisational, personal and professional concerns. These concerns do not simply dissolve in the face of the e-Science vision. Based upon our observations of healthcare professionals' data sharing practices, we argue that the success of technologies like the Grid will depend heavily on finding practical ways by which interests, rights and obligations incumbent on collaborating organisations and persons can be recognised and embedded within these technologies.

Introduction

A new generation of research infrastructure is being developed and deployed, most notably under the rubric of e-Science but also as part of national public services initiatives such as e-Health. While each domain has its own requirements, visions and problematics, the vision of transforming practice through the transparent sharing of data and computational resources is common. We are involved in several e-Science projects which plan to use the Grid to share data, including NTRAC¹, eDiaMoND² and Neurogrid³. It is premature, however, to report on data sharing practices within these projects or within e-Science generally. Instead, in this paper, we seek to draw lessons from our earlier studies of data sharing within healthcare settings and consider what their implications might be for the realisation of e-Science's ambitions.

The e-Health and e-Science visions raise a number of problems and concerns in relation to confidentiality, security, integrity, ownership, IPR, access rights and liability issues. Both struggle, on the one hand, with the necessities of preserving confidentiality of data in accord with statutory requirements and, on the other, with providing a resource for multi-disciplinary research. Systems have to be devised that meet the obligation of allowing patients to withdraw from clinical trials and, at the same time, making research data available for audit and scrutiny. Studies involving the analysis of genetic material may reveal individuals' predispositions to disease and there is an ethical obligation to make this available for purposes of healthcare provision. The relationship between e-Health and e-Science is increasingly complex as translational research begins to improve linkages between clinical practice, epidemiological studies, disease aetiology, drug development, clinical trials and new

¹ www.ntrac.org.uk

² www.ediamond.ox.ac.uk

³ www.neurogrid.ox.ac.uk

treatments. Nor is the sharing of data within and between scientific disciplines unproblematic. Researchers may not wish to contribute their hard won data to the ‘commons’ until they have published, whilst at the same time linking their data to that commons as part of their research.

Although the sharing of data for professional purposes is almost always formally governed by legal, organisational and professional rules and obligations, our studies suggest that these are played out by professionals’ practical actions in what we refer to as ‘etiquettes of data sharing’. By this, we mean the practical procedures and exigencies by which people or organisations can be seen properly and accountably to share data with others, and the circumstances under which they would be willing to do so. We are particularly interested in how sharing data depends of situated practical (professional) judgements and how organisational, legal and professional rules are worked-out in practice.

In this paper, we draw upon ethnographic investigations of a number of healthcare settings to show how healthcare professionals do not passively rely on systems (whether paper-based or electronic) to afford an appropriate degree of protection for the data that might be committed to them. Instead, we find that they make situated judgements about the appropriateness of committing data that turn upon the possible consequences of so doing informed by their understandings about how the system might subsequently restrict or allow others to access that data. The decision is not always about simply withholding or committing data but often concerns the artful selection of a particular modality or means of documenting or communicating as a practical satisfaction of the (sometimes conflicting) requirements to share data and preserve confidences.

Our contention is that judgements made as to what constitutes appropriate conduct in the use of sensitive data are profoundly situated ones, that adherence to codified ‘rules’ governing conduct is a practical problem for members and what it means to behave in accordance with those rules is something that has to be worked out afresh with each contingently arising circumstance. This is something we have referred to elsewhere as ‘practical ethical action’ (Jirotko et al., in press). We illustrate and build upon this point using examples from our fieldwork. We then explore how to approach developing data sharing technologies in ways that support engendering trust on the part of users and conclude by outlining ideas around transparency and enforcement of security policies. We present these ideas as a means of highlighting concerns that may be relevant to e-Science in general. The domain of healthcare is one where issues such as data sharing and confidentiality have already received a great deal of attention and it is our hope that, in being aware of these debates, the e-science community might find some issues with which to engage.

Confidentiality and healthcare records

Our first examples come from an ethnographic study conducted over a three year period in a unit of a UK hospital treating emergency admissions for self-harm (Hartswood et al., 2003). The unit comprised of a nine bed ward with a centrally located nurses’ station and an adjoining ‘doctors’ room’. A morning meeting following the toxicology consultant’s ward round was held in the doctors’ room, typically comprising a medical and nursing handover, followed by the allocation of patients to members of the psychiatric medical and nursing team. Before interviewing patients, team members would consult records of previous admissions (if any) to the ward. Subsequently, team members might discuss appropriate ‘disposals’ with colleagues, seek a ‘collateral history’ from others involved in the patient’s care, or from relatives, and make necessary arrangements.

Preserving confidences in talk

Of particular interest for the purposes of this paper was the orientation of the psychiatric assessment team to the patient interview, the patient's medical condition and the circumstances of their admissions as confidential matters. While not directly concerned with record keeping practice, it is instructive to examine the psychiatric assessment team's artful use of the spaces available to them to preserve various sorts of confidences. It was noteworthy that the doctors' room was treated as an area 'out of bounds' to patients and patients' relatives – patients seeking to gain the attention of the team by knocking on the door were invariably refused entry. Particular attention was given to ensuring that the door of the doctors' room was closed during discussions concerning patients and staff members neglectful of closing the door were reminded of the importance of this.

The psychiatric team were dissatisfied with the arrangements for interviewing patients. The only places close at hand where interviews could be conducted in reasonable privacy were the liaison nurse's office and the ward 'day room', both adjacent to the ward but also used for other purposes. It was typical for interviews to be conducted at the bedside (with the curtain drawn), except in circumstances where it was thought that a greater degree of privacy would be required. On one occasion, the psychiatric liaison nurse was annoyed with nursing staff for allowing a patient's relative to remain by the bedside while she was interviewing another patient in an adjacent bed. On another occasion, a doctor was reprimanded by a psychiatric assessment team member for discussing what were considered to be highly sensitive test results on the telephone at the nurses' station rather than in the doctors' room.

While it is necessary for team members to hold certain sorts of conversations with patients and colleagues, they observe particular etiquettes concerning where these conversations might appropriately be held, taking in to account their intended, as well as possibly unintended, audiences. Maintaining patient confidentiality is a practical problem for members of the psychiatric team. Working in less than ideal circumstances, they have to make judgements as to whether a bedside interview is appropriate, or if a more private setting should be sought. Team members meet their professional obligations to maintain confidences by drawing upon their mundane competences and understandings of, for example, circumstances where it might be possible to be overheard and by actively engaging in arranging matters so this should not happen, for example, by monitoring whether the door to the doctors' room is closed. The production of spaces in this way, where obligations of confidentiality can be met, cannot be separated from the routines and rhythms of the work of the ward undertaken by the various medical, nursing and psychiatric staff. These can be a source of troubles when they are misaligned. The frequent traffic in and out of the doctors' room, allowing relatives in the ward at the wrong time, or a medic discussing sensitive matters on the ward telephone leads team members to engage in various forms of alignment, maintenance and repair work – administering a reprimand, apologising to a patient, closing doors, reminding one another about policy, speaking in a quiet voice and so on. In this way, we see that the production of confidentiality as an actively collaborative endeavour and also one that has a locally accountability – because the conduct of staff is visible, both to colleagues and (to a certain extent) to patients, it is also available for scrutiny. On the one hand, the physical layout of the ward, adjoining rooms and the routine activities performed therein was a source of troubles for the team while, on the other hand, they also furnished the team with the resources for preserving confidences. Confidentiality was an on-going *achievement* by the team, dependent upon maintaining a complex choreography of conduct and location, rather than an intrinsic *property* of a place, record or utterance.

Preserving confidences in paper records

We now move on to consider the team's conduct in relation to medical records. Here again, we find that the physical organisation of artefacts and work, work's routines and rhythms, and members' competencies and understanding of these all play a part in achieving confidentiality. The first example concerns the documentation of a self-harm incident that was related to a patient's sexuality. After interviewing the patient (and following a telephone discussion with the patient's counsellor), the consultant psychiatrist proceeded to dictate two letters. The first was the routine discharge summary letter, copies of which are sent to the patient's GP, placed with the files kept on the ward and also in the patient's hospital record. The second was a letter addressed directly to the patient's counsellor. In the 'GP letter', details of the self-harm episode were sketched, omitting discussion of the patient's sexuality, a complete treatment of which was given in the letter to the patient's counsellor. When he had finished dictating, the consultant said "That should give him some protection". The consultant affirmed that his intention was to ensure discussion of the patient's sexuality was not recorded in the hospital record – but only in his correspondence with the patient's counsellor.

Protecting confidences in relation to this patient depended upon the consultant psychiatrist's familiarity with the routine ways in which discharge letters are incorporated into record corpuses, his understandings of how they then circulated and who may subsequently have access to them. His action demonstrates something of his understanding of the character and affordances of the paper medical record, that, while it should only be examined by those under same obligations of confidence, he would have effectively little control over who the subsequent audience of the record might be and under what circumstances it might be examined (the consultant has presumably judged that the details omitted from the hospital record would unlikely to have a bearing on the patient's subsequent medical care). Here we see the consultant psychiatrist's selective use of the available record keeping and communication systems⁴ in a way that artfully balances his duty to communicate the patient's admission to relevant others while, at the same time (and for all practical purposes), limiting the circulation of certain details pertaining to the case. In ways similar to the psychiatric assessment teams' management of spoken communications, confidentiality with respect to written records is also an *achievement*, depending (in this case) on the consultant psychiatrist's understandings and artful use of the affordances of the means of communication and documentation available.

Preserving confidences in electronic records

We now turn from the toxicology ward (where record-keeping was predominantly paper-based) to the work of Community Mental Health Teams (CMHTs) where a shared electronic medical record was in the process of being deployed. We have pointed previously to the role of informality in data sharing and how talk and paper records afforded provisional formulations as clinical judgements were worked up in a consensual team approach to decision-making (Hardstone et al., 2004). Because the electronic record provided for a single, canonical, non-updatable document of decisions and activities, its use in supporting the collaborative working-up of judgements was limited – rather the system worked as a repository of finalised, collectively agreed decisions. Also of interest were the etiquettes of data sharing demonstrated in the differential use of talk, paper records and computer-based records. These existed in a rough kind of hierarchy, whereby what was communicated or documented was seen as having an increasingly formal and permanent character. Thus, talk

⁴ In this case, the various letters served both to document and communicate.

afforded expressions of uncertainty, debate about alternative approaches and informal exchanges of information about clients. While provisional judgements could be written in pencil on a record template and then erased, entries into the electronic record were final. Again, members of the CMHT team oriented to the different implicativeness of documenting or communicating in these different ways with respect to the confidences of their clients. The following field work extract gives an indication of how these healthcare professionals orient to electronic medical records:

One of the CMHT team's social workers still maintained an office in the social work department (in another building, but situated close by). He had access both to the social work computer records system, as well as the CMHT system. The fieldworker had accompanied the social worker to the Social Work department to see a demonstration of the Social Work system. The social worker explained that he maintained different two sets of records, paper-based ones, and brief details of his visits on the Social Work system. He said that because any social worker has the ability to look at the electronic records he would record more sensitive information in the paper record, giving the example of a diagnosis of schizophrenia as one such example. . The fieldworker asks if the use of the system is audited – the social worker was not aware that this might be possible.

This example raises a number of important issues. First, we can see, as in earlier examples, the social worker's approach to confidentiality turns upon understandings of how access (by others) to the different records is effected. Locally held paper records are seen as a more appropriate place for 'sensitive' data where he is more confident that access can be restricted. The local visibility arrangements for a paper-based recording system, whilst informal, afford this sort of control. Although access to the electronic record may be audited (something that the social worker was surprised to learn), this is something that is not visible to him – neither the audit trail itself, nor any auditing activity. For this user, the electronic record (given what he understands about its management) provides a useful means of 'publishing' basic details about her involvement with clients to orient colleagues to the case should they have to take action out-of-hours or when the social worker is away, but not as a suitable vehicle for committing the entirety of the client's record.

This example serves to demonstrate the sorts of situated ethical judgements (Jirotko et al., in press) made by healthcare professionals as to who might access records, for what purpose, and how access can be informally supervised. Although access to electronic medical records is typically audited electronically and systems of authorisation used to restrict access, it is difficult both for audit trails to reflect the purpose of access and for authorisation mechanisms to reflect the fine-grained judgements of healthcare professionals as to its appropriateness. While it is relatively easy for access for such purposes to be managed locally for paper record corpuses, this can prove to be a cumbersome process for an electronic record system.

Taking up again the orientation of CMHT members' to preserving confidences, the following extract focuses on a discussion about the CMHT record system, specifically how access to patient records is granted in what is understood to be a shared database:

When returning with the social worker to the CMHT building, we meet with one of the team's CPNs with whom the social worker strikes up a conversation about confidentiality. The social worker asks which patients' records can be accessed on the CMHT system and by whom. The CPN is uncertain, saying that she knows that she cannot log-on outside of their team area, as has tried this when working remotely, but the system wouldn't let her on. However, she thought that she could get read access to any patient. The social worker expressed concern that this 'availability' of patient information was not consonant with the agreement that they ask the patient to sign which only permits sharing of patient data with the local team.

It is interesting that the social worker and the CPN are uncertain about patient record access controls as this is an important matter. We have seen in prior examples how it is users' understandings of how data might be subsequently accessed that informs the way such

databases are used. It is also interesting that the CPN draws upon her practical experiences of using the database to help resolve the problem. While a description of the access control mechanism is something that the CPN and social worker have either not been informed about or something that has been forgotten, it is also something that the system itself does not render transparent. It is clearly important that practitioners have an accurate understanding of access control mechanisms in order to make effective judgements concerning their management of data flows. At the point of data entry, there is nothing to tell the user what the default access policy for data entered might be nor is there any indication of the basis for that policy, the circumstances in which it is likely to hold, when it might subsequently be updated, whether there are any caveats that the user can place on use of the data or suggestions as what might not be appropriate data to record. Maintaining practitioners' awareness of the potential implications of their use of the system becomes increasingly important as the complexity of relations between data sharing entities increases but, at the same time, this may well become more difficult to achieve, as well as perhaps requiring a greater 'technical' understanding of the systems in question.

Confidentiality and e-Science

We turn now to the question of confidentiality in e-Science. The use of personal data in research is governed by a number of rules about anonymisation and informed consent and controlled by various research governance institutions such as Research Ethics Committees or Privacy Advisory Committees. While some kinds of research only require the use of anonymised or aggregated data, others require the use of identifiable personal data. Where use is made of such data, the notion of consent becomes important. The ideal situation is that fully informed people explicitly consent to the use of their data in research. This may be simple to achieve in the case of small-scale studies where people can be directly approached. In other cases, such as large-scale population-based studies, gathering explicit informed consent may not be possible. Also, some study participants may be too ill to appreciate fully the nature of the study and the ways in which it will make use of their data. Questions also arise with regard to the repeated use of data for a number of purposes – should participants be asked to re-consent each time or can they give 'blanket consent'? While participants have the right to withdraw consent at any time, how can we make sure that they are fully aware of this fact perhaps years after they have entered a study? How can we do this without unduly impacting on their lives? Sending reminders may not be appropriate and, in itself, leads to confidentiality problems. These matters are currently much debated as researchers grapple with the practicalities of research governance and data protection legislation.

Researchers in medical sciences are increasingly looking to make routine use of healthcare data⁵ and, in some instances, there may be a flow of data in the other direction. This means that those handling data in either context need to be aware of the potential uses to which it may be put. For example, letters may need to be written to patients, to their GPs, to other carers to inform them about matters relating to their participation in a study or it may be crucial for a person's routine or emergency care for healthcare professionals to be aware that they are taking part in a trial. These examples emphasise that clinical research and clinical practice are impossible to separate analytically or practically. While research use of personal data is subject to formalised and regulated controls, we argue that research staff will also need to engage in the same kinds of practical ethical action described in this paper. Though we have not studied this as extensively as healthcare, we can present two illustrative examples that highlight the issue.

⁵ See, for example, the CLEF project. www.cs.man.ac.uk/mig/projects/current/clef

A study collecting food frequency data from participants was facing the problem that people often cannot fill in detailed questions about their food intake without referring to their fridges or their partners. It was thus decided that participants should be given the questionnaire to fill in at home rather than in a clinic. This raises the issue of disclosure of data through the interactions necessary to administer this process. Clearly, a questionnaire needs to contain identifiers so it can be linked to the correct record but participants often feel uncomfortable about sending such data by post and would sometimes remove the relevant part of the questionnaire, effectively rendering the data useless.

Research nurses recruiting study participants have to collect various kinds of data, e.g., basic demographics needed for record linkage, baseline phenotype data, medical history or family history. This work may take place in special clinics or in healthcare settings such as at the bedside or at follow-up clinics. Research nurses thus face issues similar to those of the psychiatric assessment team. They may face additional difficulties when they are doing this other than in their normal workplaces where they may not be familiar with local arrangements. These issues lead to a potential for conflict between carers (who may play the role of participant gatekeepers) and research staff.

Discussion

Our investigations of data sharing have focused on the situated practices, artefacts and knowledge underpinning the conduct of healthcare delivery and scientific practice. Common to these settings are a variety of data bearing artefacts on a range of different media – too numerous to list exhaustively – but comprising a heterogeneous array of reports, records, notebooks, computer systems and media, including artefacts specific to disciplines such as X-Ray films, gels, blots, printouts, lab-books, medical records and so on. Some have a formal character, whereas others are more informal; some have a greater degree of permanence, others are more temporary; certain sorts of representations may reoccur, others may be created to solve a particular problem at hand. These artefacts occupy places in complex ecologies of practice: they occupy physical spaces – there are places where they can be routinely or expectably be found, they are organised and ordered in predictable ways. They both constitute and are constituted by the unfolding trajectory of the work at hand as they are shared, moved between different work locations, become the responsibility of different people and the locus of different sorts of work as well as being a means of coordinating work between different people. They are transformed in various ways: edited, annotated, amalgamated, copied, archived and discarded. We can extend this sketch and consider how the life-world of workplace artefacts intersects with that of talk: in different situations, to different audiences, at different times and in different places presents as rich a variation in purpose and organisation and is deeply intertwined with practice as are artefacts.

Although delineating the different orders that talk and artefacts constitute, play a role in or belong to (various spatial, temporal, praxiological orders), the above sketch neglects their relation to the multiple and intertwined *moral orders* of the workplace. The first of these we can identify as being associated with the conduct of the work itself – the various rights, obligations and expectations that are constitutive of the working division of labour within a workplace. These comprise members' shared (but contestable) understandings about whose job it is to do what and when, where members' responsibilities begin and end, how they might be appropriately discharged, transferred or delegated. A second, related moral order, and one central to this paper, concerns the responsibilities associated with membership of a profession and of an organisation – that conduct should be in line with the various rubrics of behaviour associated with these incumbencies, typically codified as organisational rules,

professional codes and legal obligations. There are a wide range of issues here, spanning a number of domains but including confidentiality, security, research ethics and scientific accountability.

Confidentiality and media affordances

Discussions concerning the merits and demerits of records illustrate the chimerical character of media as both a source of troubles and as a resource with respect to maintaining confidences. For example, the Electronic Record Development and Implementation Programme (ERDIP) study report into patient consent and confidentiality argues that paper record systems serendipitously afford restricted circulation and sharing (ERDIP, 2002). In fact, the examples from our fieldwork show that decisions as what to record are carefully made in light of the medium's affordances. Paper medical records can afford a rough and ready access control mechanism as part of their physical organisation. A contrasting view can be found in the Confidentiality and Security Advisory Group for Scotland (CSAGS) report to Scottish ministers: "Present records are maintained on paper. This makes it very difficult to restrict the access of different professionals to the information that they need to know" (CSAGS, 2002).

Similarly, it is claimed that integrated care records pose a threat to confidentiality, principally because records become massively available (i.e., available to many potential readers at remote sites):

"Automation and, more importantly, networking have changed this situation radically. Data have no physical embodiment, are easily copied, and are accessible from multiple points of access. Large numbers of records can be transferred as easily as a single one. The existence of the Internet means that data can be moved across administrative, legal and national jurisdictions as easily as it can be moved to the next desk; intrusions can be mounted with equal felicity. Electronic medical records also raise the possibility that much more accurate and complete composite pictures of individuals can be drawn – so much more so that reasonable people would raise concerns about the aggregate even if they had no concerns about any single data element" (Committee on Maintaining Privacy and Security in Healthcare Applications, 1997).

Also, and at the same time, they can be more 'secure' than paper records, because access to them can be properly audited and access controlled to ensure that it is more strictly realised on a 'need to know' basis (Barrows and Clayton, 1996). It can be argued that these 'properties' of record-keeping systems are little more than *potentialities* dependent on the specific details of a particular instantiation. Although it is not uncommon for paper medical records to be judged in the light of their electronic counterparts (and vice versa), perhaps like is not always compared with like. True enough, paper records do not necessarily afford the sorts of fine-grained access controls that can be extended to electronic records but, on the other hand, because of their physical character – they can only be in one place at one time – those accessing the records are visible to their colleagues, a sort of local informal accountability is afforded that may be 'good enough'. The requirement for stricter audit and access controls of electronic records derives from their wider, non-local availability. Not only is it easier to enforce fine grained access control and audit mechanisms in electronic medical records, it is arguably more important for this mode of record keeping that such mechanisms are in place. Moreover, if carefully designed, electronic media have the potential to provide a variety of security affordances that mesh more closely with the ethical practices of users. One feature of our field work is the way that healthcare professionals were seen to make choices between different record keeping modalities and systems available to them contingent on the ethical and practical concerns at hand. Provision of a 'one size fits all' security policy does not make sense in this context nor, perhaps, does trying to anticipate all of circumstances

which a user might face. It might be more appropriate to think about how to provide tools for users to manage the circulation of data they enter onto the system.

It is perhaps unsurprising that integrated care record systems have proved to be controversial, not least for reasons relating to confidentiality (Oldfield, 2003). One might view these concerns (without denigrating them) as in part arising from the *transition* itself. We have seen how healthcare professionals based their judgements on what data should be committed and how on their understandings of the routine ways and circumstances under which the record might be subsequently accessed. This turns, in part, on a biographical familiarity with the system(s) in question – the methods that have been used, or been seen to work in the past. The existing system comes with the guarantor that, despite shortcomings, it seems to have worked in the past, is understood and, by virtue of these understandings, affords a degree of control to the professionals using it. On the other hand, in the introduction of a system on the scale of the UK National Programme for IT (NPFIT), not only are the risks emphasised but also there is very little practical understanding of the likely consequences of committing data to the record. Thus, NPFIT may be seen only as a source of potential troubles with respect to maintaining confidentiality – healthcare professionals lacking the necessary practical engagement with the system to see how it might also be a resource for this purpose.

Locating policy decisions and control

Where e-Health (and by implication, e-Science) applications are devised to support clinical practice, we find that the confidentiality policies become tightly embedded in the system and centrally administered, disallowing the sorts of routine situated ethical judgements (about, for example, who might appropriately view or edit a record) observed in our studies.

Legislation and regulatory frameworks shift of responsibility for privacy and confidentiality centrally. Technology enables more centralised control. Matters of confidentiality handled routinely and locally by professionals with respect to paper records can more easily be codified, formalised and inscribed in computer systems. More precisely, the extent to which data available is located in institutionally recognised and sanctioned media (computer systems, paper documents, portable electronic media), and transmitted in institutionally recognised and sanctioned ways (fax, email and telephone conversations), its transmission, organisation and access can be subject to organisational policy, regulation and control. However, around the margins there may be a whole host of data bearing artefacts and communication practices that do not form part of the ‘official’ organisationally sanctioned and accountable corpus. For example, ad hoc notes taken by nurses on scraps of paper (Hardy et al., 2000), notebooks, diaries, post-it notes, ‘off-the-record’ conversations and so on.

The centralising of policy and its enactment (becoming an organisational rather than a professional matter) shifts the responsibility for ensuring confidentiality and, concomitantly, the risks of disclosure (or otherwise breaching the law) to some management or administrative post, divorced from the concerns of the settings where data is gathered and consumed. Such posts may be mandated by legislation, for example, Data Protection Officers and Cauldicott Guardians. This can lead to a conservative approach to privacy and confidentiality. For example, the ERDIP report (2002) states that a number of ERDIP projects have been blocked because “Individual gatekeepers are concerned about their roles, responsibilities and personal liability if they should make the wrong choice ... it is generally easier for them to say that data cannot be shared (for whatever reason seems most appropriate) rather than to authorise the process.”) in a way that is at odds with creativity and innovation, and challenging to the status of healthcare workers as professionals.

One solution is to find ways that control can be delegated to ‘front-line’ professionals, particularly in non-routine contexts of access described above. Some progress has been made in this direction in e-Science projects (e.g., Power, 2005). Mechanisms that allow delegation, while not being free from risk, do have the advantage of ensuring a more complete audit trail, and provide a means to ensure activities remain organisationally accountable (rather than, for example, being conducted in ad-hoc ways, through the use of another’s password or on a locally held and separately maintained records system).

Conclusions and recommendations

Our observations of healthcare professionals’ data sharing practices reveal that they routinely make fine-grained judgements concerning with whom it is appropriate to share data and when. They demonstrate artfulness in their choice of different communication and record keeping modalities that turns upon their understandings of how far recorded or communicated data might circulate, and to whom. We argue that it is through their mundane and ‘biographical’ familiarity with the mechanics of record handling and distribution that healthcare professionals can come to trust those systems; more precisely, to know when it is appropriate to trust particular systems and when it is not. Trust depends upon users’ practical understandings of what that system affords in the way of preserving confidences, which (even for relatively uncomplicated systems) can sometimes be uncertain or inaccurate. Where more complex systems are envisaged that either support data sharing between organisations to enhance service provision or data linking to support research, then it may become increasingly difficult for users to make judgements about the consequences of committing data to the system. We find that the integrity of confidences depends not only on security mechanisms but also on the artful use of those affordances as part of users’ situated judgements as to *when* to commit *which* data to *what* system. It makes sense, therefore, to pay close attention in design to providing the sorts of affordances that more closely match the requirements of practice. Otherwise, the risk remains that alternative means of communicating and documenting data may be improvised and used preferentially.

We argue that our findings have a number of implications for e-Health and e-Science. With the increasing focus on using administratively derived data for policy and research purposes⁶, questions arise as to the quality and completeness of potential data sources. It is important to have an understanding of the practices of those generating the data to generate an awareness of the sorts of data that might routinely be omitted and also to find ways of engendering trust to minimise such omissions. It is beyond the scope of a single paper, however, to provide detailed recommendations for addressing the issues raised or even to give each individually the detailed treatment it deserves. Rather, our aim has been to draw some preliminary conclusions that might help those engaged in working with or assembling Grid infrastructures and tools:

Control: the central issue is control over data – who decides access, who has access and for what purposes? e-Science may involve individual researchers surrendering some control over others’ access to their data and the purposes to which it might be put. Data sharing in e-Science requires, on the one hand, recognition of the importance of the ‘commons’ of data and, on the other, an understanding of what it means to allow data ‘into the wild’. Crucially, data ‘in the wild’ does not inevitably entail the kinds of rights and obligations set out by Mauss (1954) in his discussion of ‘gift-giving’ – there is no in principle reciprocity involved

⁶ In the UK, for example, the use of administratively derived data seems likely to replace traditional methods such as the Census.

in data sharing in e-Science (although this may exist in practice in the form of acknowledgements and so on).

Trust: a related issue is ‘trust’ – how can one know that data will be used appropriately when it is ‘in the wild’? As noted above, there is no in-principle reason why some uses of data should be closed off and, in practice, it is difficult to police such uses once the data have become part of the ‘commons’. One solution would be for those who generate data to ‘recipient design’ datasets so that potential users will be aware of the intentions of the originator and the uses to which data might be put (i.e., its provenance). Of course, one cannot know how widely data will be circulated (to whom and with what aim) but recipient design would provide some form of indication as to what is a warrantable use of the data.

Transparency: Committing data to shared databases requires an understanding of the access policies that apply and their implications, particularly where complex relationships exist between several databases. How does e-Science make sense of access policies in shared databases and how can those policies be made transparent? Confidentiality is an achievement on the part of users, i.e., judgements as to whether it is appropriate to commit data depend upon understandings of what might subsequently happen to that data. These may be difficult to come by for even moderately complex systems and it may be useful to have a means to query the system as to what the implicativeness of committing data to the system might be.

Accountability: Means of enforcing access policy (e.g., audit trails, authorisation, authentication etc) can be seen as ways of repairing accountability where data can be accessed independent of location and by a large user population. How can accountability be maintained in large distributed systems? Although users might be reassured that access logs are adequately policed, this activity tends to be opaque unless an actual breach is identified and publicised. Making audit trails visible, not only to those who have a specific responsibility for audit but also to those who are generators of the data, may help provide a surrogate for visibility arrangements that hold for locally held records, where access to data is open to scrutiny and challenge. This would also assist users in developing a biography of the system and its affordances, i.e., a practical understanding of the consequentiality of committing data.

Etiquettes of data sharing suggest that data is not an unencumbered commodity that can simply be shared on a ‘no strings’ attached basis. Data becomes entangled with obligations at the moment of its telling or sharing. This involves a nexus of rights, obligations and expectations associated with (*inter alia*) collegiality, data ownership, data protection, IPR, competitive advantage (both in the commercial and academic worlds), ethics, and other organisational, personal and professional concerns. These concerns do not simply dissolve in the face of the e-Science vision. Rather (as we have shown above), they are thrown into sharp relief as the recognisable and well understood affordances of physical spaces and local data storage systems, and the knowledge of organisational practices and rules – i.e., those things upon which people observably rely when making professionally accountable judgements about sharing data – are seen against a backdrop of the non-physical spaces, global infrastructures and virtual organisations that are characteristic of the e-Science vision. The success of data sharing technologies such as the Grid will depend heavily on finding practical ways by which various sorts of interests, rights and obligations incumbent on collaborating persons and organisations can be recognised and embedded within them.

Acknowledgments

We wish to thank all the healthcare practitioners who participated in the studies for their time and patience. This work was funded by the UK EPSRC (grant numbers GR/R86751/01 and GR/M52786), the UK ESRC/EPSRC Dependability Interdisciplinary Research Collaboration (DIRC), and was also supported by the Scottish School of Primary Care and Stirling University.

References

- Barrows Jr, R. C., Clayton, P. D. (1996). Privacy, Confidentiality and Electronic Medical Records. *Journal of the American Medical Informatics Association*. 3:2 139-148.
- Committee on Maintaining Privacy and Security in Health Care Applications of the National Information Infrastructure (1997). For the record – Protecting Electronic Health Information. National Academy Press. Washington DC. Online version available at: http://www7.nationalacademies.org/cstb/pub_fortherecord.html
- The Confidentiality and Security Advisory Group for Scotland (CSAGS) (2002). Final report. Protecting Patient Confidentiality. www.show.scot.nhs.uk/sehd/publications/ppcr/ppcr.pdf
- Electronic Record Development and Implementation Programme (ERDIP) (2002). Study report N5–Patient consent and confidentiality. Cambridge Health Informatics Ltd. www.nhsia.nhs.uk/erdip/pages/evaluation/docs/consentconfid/Consentstudyreport.pdf.
- Hardey, M., Payne, S. and Coleman, P. (2000). ‘Scraps’: Hidden nursing information and its influence on the delivery of care. *Journal of Advanced Nursing*, 32 (1) 208-214.
- Hardstone, G., Hartswood, M., Procter, R., Rees, G., Slack, R. and Voss, A. (2004). Supporting informality: Team working and Integrated Care Records. In Proceedings of the ACM Conference on Computer-Supported Cooperative Work, November.
- Hartswood, M., Procter, R., Rouncefield, M. and Slack, R. (2003). Making a Case in Medical Work: Implications for the Electronic Medical Record. *Journal of Computer Supported Cooperative Work*, 12, pages 241-266.
- Jirtoka, M., Procter, R., Hartswood, M., Slack, R., Coopmans, C., Hinds, C. and Voss, A. (to appear). Collaboration and Trust in Healthcare Innovation: the eDiaMoND Case Study. *Journal of Computer-Supported Cooperative Work*.
- Mauss, M. (1954). *The Gift: The Form and Reason for Exchange in Archaic Societies*. Glenco. IL: Free Press.
- Oldfield, P. D. (2003). Sharing patient information electronically throughout NHS: Patient confidentiality may not be guaranteed. *BMJ*. 327:623.
- Power, D., Slaymaker, M., Politou, E. and Simpson, A. (2005). A secure wrapper for OGSA-DAI. In Proceedings of the European Grid Conference. February 14-16, Science Park Amsterdam, The Netherlands.