

Executive summary

Key findings and recommendations

1. There is a growing accumulation of data, of increasing variety, about human biology, health, disease and functioning, derived ultimately from the study of people. Advances in information technology and data science provide more ways, and more powerful ways, to collect, manage, combine, analyse and derive insight from these data. The result is that data are now seen as a valuable resource with an indefinite range of potential uses.
2. There is a public interest in the responsible use of data to support advances in scientific knowledge, innovative treatments and improvements in health services. However, there is also a public interest in protecting the privacy of individuals: privacy is fundamentally important to individuals (and groups) in the establishment and maintenance of their identity, their relationships and their sense of personal well-being. In biomedical research and health care data initiatives, which link and re-use data, public and private interests are entangled in complex ways. Such data initiatives must address the following question:
 - what is the set of morally reasonable expectations about the use of data and what conditions are required to give sufficient confidence that those expectations will be satisfied?
3. Compliance with the law cannot guarantee that a use of data is morally acceptable. Faced with contemporary data science and the richness of the data environment, protection of privacy cannot reliably be secured merely by anonymisation of data or by using data in accordance with the consent from 'data subjects'. Effective governance of the use of data is indispensable.
4. A set of morally reasonable expectations about the governance and use of data should be determined in accordance with four principles:
 - the principle of respect for persons
 - the principle of respect for established human rights
 - the principle of participation of those with morally relevant interests
 - the principle of accounting for decisions
5. Taking into account the current state of knowledge and practice, and the likely direction and pace of developments, and considering a number of specific data initiatives in biomedical research and health care, we recommend:
 - support for needed research into the potential harms associated with abuse of biological and health data, as well as the benefits of responsible data use
 - comprehensive mapping of UK health and research data use and the norms relevant to it
 - mandatory reporting of privacy breaches affecting individuals to the individuals affected
 - review of anti-blagging measures to protect data in health care systems and promulgation of best practice
 - criminal penalties, including imprisonment, for the deliberate misuse of data

- a public statement of expectations about who may be given access to health data and for what purposes, for each data initiative
- publication of all Health and Social Care Information Centre data sharing agreements and results of independent audits of compliance
- maintenance of an auditable record of all people given access to data held by the Health and Social Care Information Centre, that can be given to affected individuals
- a review of the appropriateness of public-private partnerships to secure public benefit from the research use of National Health Service records
- increased subject participation in design and governance of research projects
- wider use of explicit and flexible ethics and governance frameworks for research projects, including for international collaborative research
- restriction of access to research data to researchers (including international collaborators) who are subject to institutional oversight and effective sanction
- publication of policies on the use of cloud services by national bodies
- ethical and scientific appraisal to maximise the contribution of participant-led research to science while ensuring adequate protection of participants
- collaboration among all members of the research community to promote a more robust, explicit and candid foundation for extending access to data for research in the public interest.

Summary of the report

6. This report takes as its starting point the massive accumulation of data in biomedical research and health care, and the increasing power of data science to extract value by linking and re-using that data, for example in further health or population research. It examines the scientific, policy and economic drivers to exploit these opportunities, and the concerns and potential risks associated with doing so. The faltering ability of conventional information governance measures to keep pace with these developments is identified as a significant problem. The report therefore poses and addresses the following question:
 - how can we define a set of morally reasonable expectations about the use of data in any given data initiative and what conditions are required to give sufficient confidence that those expectations will be satisfied?
7. The report sets out a number of general recommendations, including four guiding principles for ethical design and governance of data initiatives. These help to identify specific examples of existing good practice and to make recommendations for improved practice in the use of data in the fields of health care (re-use of NHS records, clinical research, etc.) and population research (biobanks, epidemiological studies, etc.).

Data (chapter 1)

8. Data provide the raw materials for reasoning and calculation. The informational value of data arises from the context in which they are placed, and how they relate to other data. The meaning, utility and value of data may be transformed as they appear within different contexts such as health care, research and public policy. Digitisation has allowed an escalating accumulation of data in health care and biomedical research settings, including:
 - clinical care data (e.g. primary care and hospital records)
 - data from clinical trials and observational studies
 - patient-generated data (e.g. from 'life logging' or consumer genetic testing)
 - laboratory data (e.g. from imaging, genome sequencing and other 'omics')

- administrative data or metadata
9. Advances in information technology (faster information storage, retrieval and processing) and data science (more powerful statistical techniques and algorithms) have created novel opportunities to derive insights from the analysis of big datasets, and particularly through the combination or linking of datasets. While these developments are not specific to biomedical research and health care, they are having a significant impact in these fields, with morally significant implications. They have led to the emergence of a new attitude towards data that sees them as exploitable raw materials, which can be put to use for a variety of purposes beyond those for which they were originally collected.
 10. ‘Data initiatives’ involve the re-use of data in novel contexts and linking them with data from other sources. However, inconsistent data quality and peculiarities arising from the context of data collection can present technical difficulties in exploiting these opportunities. Furthermore, legal and ethical limitations placed on the re-use of data for secondary purposes limit the re-use of existing data sets.

Opportunities and threats (chapter 2)

11. The combination of advances in information technologies and in data science have generated considerable opportunities for economic activity. Given the UK’s strong research base in the biomedical sciences and the unique resource and infrastructure of the UK’s national health services, the use of health data has become a strategic focus.
12. There is a clear public interest in the responsible use of data to improve well-being through improved health advice, treatment and care, as well as through increasing economic prosperity more generally. This objective is being pursued in three main ways:
 - increasing efficiency and transforming service delivery through better informed decisions about resource allocation and greater involvement of patients through e-health care
 - generating improvements in medical treatment by building a stronger evidence base for prediction, prevention and treatment, and by using data to personalise treatment and care, linking phenotype and genotype data with lifestyle, environmental and social data
 - generating economic growth from the life sciences by using existing data in health systems with increased technological capacity and skills to invigorate the pharmaceutical and biotechnology industries.
13. To achieve these outcomes a number of policy orientations have been set in the UK and elsewhere, such as:
 - increasing IT intensity and introducing new infrastructure in health systems
 - establishing partnerships between the public and private sectors to promote resource exploitation and innovation
 - centralising data resources to facilitate analysis of linked data
 - promoting ‘open data’ and ‘data sharing’ to encourage the widest possible use of resources
 - promoting ‘big data’ and investing in the knowledge economy to foster development of new tools, methodologies and infrastructures.

14. However, the pursuit of opportunities must take account of the need to manage a number of threats to welfare. These threats take a number of forms, for example:
 - misuse of data leading to harms to individuals and institutions (ranging from detriment to health, loss of privacy, financial loss, reputational damage, stigmatisation and psychological distress)
 - discriminatory treatment, ranging from targeted advertising to differential pricing that compounds social disadvantage, to discrimination in insurance and employment
 - state surveillance of citizens, particularly in the light of revelations about the US National Security Agency, which is greatly facilitated by large databases and linked information systems
15. Independent research commissioned to inform our work found that the negative impacts of data misuse are potentially much wider than are those recognised by legal and regulatory systems. Furthermore, the nature of privacy harms and of the judicial and regulatory systems means that they are likely to be under-reported by the victims and obtaining redress is difficult.

Recommendations

- 1 Relevant bodies, including public and private research funders and UK health departments, should ensure that there is continued research into the potential harms associated with abuse of biological and health data, as well as its benefits. This research should be sustained as available data and data technologies evolve, maintaining vigilance for new harms that may emerge. Appropriate research that challenges current policy orientations should be particularly encouraged in order to identify and test the robustness of institutional assumptions.
- 2 The Independent Information Governance Oversight Panel and the Health Research Authority should supervise, respectively, the maintenance of comprehensive maps of UK health and research data flows, and they should actively support both prospective and continuing evaluation of the risks or benefits of any policies, standards, or laws governing data used in biomedical research and health care.
- 3 The Government should make enforceable provisions to ensure that privacy breaches involving individual-level data that occur in health services and biomedical research projects are reported in a timely and appropriate fashion to the individual or individuals affected.
- 4 The Health and Social Care Information Centre should maintain prospective assessments to inform the most effective methods for preventing the inadvertent or fraudulent accessing of personal healthcare data by unauthorised individuals.
- 5 The UK government should legislate to introduce criminal penalties, comparable to those applicable for offences under the Computer Misuse Act 1990, for deliberate misuse of data whether or not it results in demonstrable harm to individuals.

Moral values and interests (chapter 3)

16. The concept of *privacy* and the distinction between public and private have evolved throughout history. Individual privacy is important in the formation of identity and the maintenance of personhood but privacy can also be attributed to families, and wider groups. Norms of information disclosure are important in the formation and maintenance of identities and relationships between individuals and groups, and different norms apply to different relationships.

17. An important class of privacy norms is enshrined in the rules and practices of *confidentiality*. These may exist as informal conventions among individuals but may be more formalised in professional relationships, contracts and laws. Medical confidentiality allows information sharing that might otherwise infringe privacy norms to take place for specific professional purposes.
18. At the same time *consent* provides a mechanism to make controlled exceptions to an existing privacy norm for specific purposes (for example, to permit a medical diagnosis or referral) without abolishing the underlying norm. However, consent does not itself ensure that all of the interests of the person giving consent are protected nor does it set aside the moral duty of care owed to that person by others who are given access to the information. On its own, consent is neither necessary nor sufficient for ethical extensions of data access.
19. While individuals have privacy interests in the use of data, they also share group interests in the wider use of data for health research. This broader *public interest*, which consists in securing objectives that are valued by society, may come into conflict with individual privacy. But the relationship between privacy and public interest in data is not simply one of opposition. The two are mutually implicated in each other: there are private interests in the achievement of common goals and a public interest in the protection of privacy that encourages cooperation. This complex relationship leads to a need to reconcile the articulation of the private within the public and the public within the private. A fundamental moral question facing data initiatives is therefore:
 - how may we define a set of morally reasonable expectations about how data will be used in a data initiative, giving proper attention to the morally relevant interests at stake?
20. Three sorts of considerations will be relevant to formulating an answer:
 - the norms of privacy and disclosure applicable among those who participate in a data initiative
 - the ways in which individual freedoms are respected, for example, the freedom to modify these norms by consent
 - the form of governance that will give acceptable assurance that the expectations will be met

Law, governance and security (chapter 4)

21. A number of overlapping legal measures exist to protect privacy, principally: formal privacy rights, which guarantee freedom from interference, albeit that they may be qualified by certain public interest considerations; rules of data protection, which control the 'processing' of various kinds of 'personal data'; and duties of confidentiality, which protect against unauthorised or unreasonable breaches of confidence.
22. A number of technical measures may be applied to prevent the identification of individual subjects and reduce the risk of privacy infringements:
 - aggregation of data makes it harder to distinguish individual cases, although it is not wholly secure in the face of modern statistical techniques; it also makes further linking of data difficult

- anonymisation by the removal of identifiers also makes individuals difficult to re-identify, although re-identification may still be possible depending on what other data or information are available
 - pseudonymisation, the replacement of identifiers with a code, enables linking of data where the correspondence between the code and the case is known, although data may still be vulnerable to inferential re-identification
23. While de-identification measures may help to protect privacy, re-identification may not be impossible and the risk of re-identification is both difficult to quantify and may become greater over time. To protect the privacy of data subjects, de-identification should therefore be combined with further controls on the access to and uses of data.
24. A standard control is to limit access to data in accordance with the consent of the 'data subject'. Broad consent allows data subjects to set certain parameters for the use of the data that are morally salient for them but the often far-reaching implications of data use may be obscure, and the scope of consent given when data are collected may become unclear, particularly in changing circumstances and in relation to novel uses of data. This is especially likely when data are held for long periods of time. Continuing involvement of subjects through 'dynamic' forms of consent can address this but is potentially demanding.
25. While seeking consent respects rights that individuals may have to make decisions about matters that may affect their interests, it cannot protect them from potentially harmful consequences of data use. Merely acting in accordance with consent cannot excuse data users from their moral duties towards data subjects, indeed towards all those who have a morally relevant interest in the data initiative, whether they are people from whom the data were initially collected or others who stand to be affected by their use.
26. As neither anonymisation nor compliance with consent offer sufficient privacy protections in data initiatives, additional controls on the use of data – on who is permitted to access them, for what purposes, and how they must conduct themselves – are therefore required. These have administrative aspects (e.g. data access committees and agreements) and technical aspects (e.g. safe havens).
27. The need to meet two contradictory requirements at the same time places data initiatives in a double bind. In other words, they are required:
- to generate, use and extend access to data (because doing so is expected to advance research and make public services more efficient); and, at the same time
 - to protect privacy as this is a similarly strong imperative, and a requirement of human rights law (and the more access is extended the greater the risks of abuse).
28. In order to meet this challenge the use of measures such as anonymisation, the mechanisms for respecting individuals' rights and interests (such as consent procedures), and the forms of governance that guide the conduct of professionals in the public interest need to be established coherently. These measures should be determined in relation to the underlying moral norms and values, and in relation to an agreed understanding of the hazards, benefits and uncertainties of data use in the context of particular data initiatives.

Ethical governance of data initiatives (chapter 5)

29. Data initiatives are practical activities that involve a number of actors (who might be individuals, groups, institutions, etc.) some of whom stand to benefit or lose from the outcomes. Tensions and potential conflicts between values and interests can arise at the level of the individual, of professions or of the public. The ethical formation of a data initiative is a matter of reconciling these values and interests in a coherent set of morally reasonable expectations.
30. A morally reasonable set of expectations should embody four principles.

Ethical principles for data initiatives

The use of data in biomedical research and health care should be in accordance with a publicly statable set of morally reasonable expectations and subject to appropriate governance.

- **The set of expectations about how data will be used in a data initiative should be grounded in the principle of respect for persons.** This includes recognition of a person's profound moral interest in controlling others' access to and disclosure of information relating to them held in circumstances they regard as confidential.
- **The set of expectations about how data will be used in a data initiative should be determined with regard to established human rights.** This will include limitations on the power of states and others to interfere with the privacy of individual citizens in the public interest (including to protect the interests of others).
- **The set of expectations about how data will be used (or re-used) in a data initiative, and the appropriate measures and procedures for ensuring that those expectations are met, should be determined with the participation of people with morally relevant interests.** This participation should involve giving and receiving public account of the reasons for establishing, conducting and participating in the initiative in a form that is accepted as reasonable by all. Where it is not feasible to engage all those with relevant interests – which will often be the case in practice – the full range of values and interests should be fairly represented.
- **A data initiative should be subject to effective systems of governance and accountability that are themselves morally justified.** This should include both structures of accountability that invoke legitimate judicial and political authority, and social accountability arising from engagement of people in a society. Maintaining effective accountability must include effective measures for communicating expectations and failures of governance, execution and control to people affected and to the society more widely.

31. The principle of respect for persons does not mean that individuals' interests may never be overridden, but that they may only be overridden where there is a legitimate reason to do so. As a principle of design of data initiatives, the principle of respect for human rights seeks to avoid potential rights conflicts and violations rather than leaving them to be dealt with retrospectively through judicial processes. The participation of people with morally relevant interests in the design and governance of data initiatives allows the identification of relevant privacy norms and the development of governance measures (such as design of consent and authorisation procedures) in relation to these norms; it allows preferences and interests to be expressed and transformed through practical

reasoning, and account to be given of how these interests are respected in decision making, helping to foster trust and cooperation. The principle of accounting for decisions ensures that expectations, as well as failures of governance and control, are communicated to people affected and to others more widely. It also ensures that data initiatives remain in touch with changing social norms.

Data initiatives in health systems (chapter 6)

32. Health-care IT systems were originally introduced to facilitate basic administrative tasks, such as managing patient records, issuing repeat prescriptions and tracking patients through different encounters with health care professionals. However, they developed to provide business intelligence for service improvement and support for observational research. These come together in the concept of a 'learning health system' which is seen as an inevitable response to increasing pressures on health services and the demand for new treatments.
33. These functions, together with the need to manage reimbursement and the appetite for data to inform health policy, combined to push data systems in the English NHS towards a centralised approach with electronic health records at its heart. The central collection of health care data in England is now managed by the Health and Social Care Information Centre (HSCIC). The HSCIC's model involves holding linked data centrally, publishing aggregate data and disclosing certain individual-level 'pseudonymised' data in controlled conditions. Debate around the 'care.data' programme to extract primary care data to the HSCIC focused attention on the assumptions made about the relationship between privacy norms relevant to NHS patients and the legal norms under which HSCIC operates. It highlighted the absence of governance arrangements to negotiate this difference, and raised questions about how the rights of individuals were respected. Failure to attend to these prospectively led to ad hoc policy changes and a damaging loss of public and professional trust.
34. The Scottish Informatics Programme to develop a research platform for health data involved initial public consultation to identify relevant social norms. On the basis of this it developed a model whereby data linkages are performed for specific purposes using a trusted third party, with analysis carried out on linked, pseudonymised datasets in a controlled environment. Data are not warehoused centrally and no individual-level data may be released from the safe haven. Direct access to the data is not available to commercial researchers. In addition to the role of the data custodians, authorisation for data use is provided by a risk-based, proportionate governance system that takes into account both privacy risk and public benefit, and refers to an explicit, potentially revisable, statement of guiding principles and best practices. The model demonstrates a number of features of good practice in relation both to its development and form that are consistent with the principles set out in this report.
35. The 100,000 Genomes project involves linking data from genome sequencing with individuals' NHS records for the investigation of cancers, rare diseases and some infectious diseases. Broad consent is obtained from individual subject participants (who do not expect direct therapeutic benefit). In operation the project will have an ethics committee and an explicit data access policy. Authorised researchers from all sectors may access a firewall-protected, pseudonymised dataset administered by a Government-owned company, Genomics England Ltd. No individual-level data may be released from this environment. The claimed public interest lies explicitly in securing economic as well as scientific and therapeutic benefits, by stimulating the commercial sector.

Recommendations

- 6 An independent, broadly representative group of participants should be convened to develop a public statement about how data held by the HSCIC should be used, to complement the Code of Practice on confidential information. This should clearly set out and justify what can reasonably be expected by those from whom data originate and be able to demonstrate that these expectations have been developed with the participation of people who have morally legitimate interests in the data held by the HSCIC, including data subjects, clinical professionals and public servants.
- 7 In addition to implementing the recommendations of Sir Nick Partridge's review, all Data Sharing Agreements held by the HSCIC should be published, along with the findings of a periodic independent audit of compliance those agreements.
- 8 HSCIC Data Sharing Agreements should include a requirement to maintain an auditable record of all individuals or other legal entities who have been given access to the data and of the purposes to which they are to be put; this should be made available to all data subjects or relevant authorities in a timely fashion on request.
- 9 Broader public consideration should be given to whether Genomics England Ltd provides the most appropriate model for the ethical use of genomic information generated in health services for public benefit before it becomes the *de facto* infrastructure for future projects.

Population research (chapter 7)

36. Biobanks are major resources of tissues and data that may be used for a variety of research purposes. They support the trend in life sciences research towards broader collaboration, larger datasets and greater varieties of data. UK Biobank is a large population biobank established to support the investigation of a range of common diseases occurring in the UK. Subject participants give broad consent to the use of data collected at recruitment, from their medical records and through supplementary data collections (e.g. the imaging study). The resource has a published Ethics and Governance Framework, compliance with which is overseen by an independent Ethics and Governance Council. Its design was foreshadowed by meaningful public engagement but the intention to establish a participant panel was not followed through. Subject participants' influence over the use of the data is limited to possible withdrawal from the resource on the basis of information published or communicated by the organisation. Pseudonymised data are released to researchers from recognised institutions for research that meets public interest criteria, and results are returned and published to support further research. There may be some need to review the set of expectations underlying the operation of UK Biobank in the light of changing circumstances (the evolving data environment, revaluation of data, etc.) One such area is feedback of information to subject participants; another is expectations about commercial access to the resource. Renewed engagement with public and participants is desirable in this context.
37. The UK10K Rare Genetic Variants in Health and Disease project was established to use existing research samples to characterise the genetic bases of rare diseases through comparison of genotypes of affected individuals with deeply phenotyped groups from cohort studies. This confronts the problem of controlled disclosure of highly specific individual-level data among different groups of researchers working on distinct studies. It is achieved through a common ethical framework of policies and guidelines developed with some patient interests represented. It places considerable reliance on ensuring that

only appropriately qualified researchers, bound by enforceable agreements, have access to data and on the role of principal investigators as data custodians.

38. International collaborative research initiatives such as the International Cancer Genome Consortium and the Psychiatric Genomics Consortium also rely on a common ethical framework operating across different research contexts. These need to accommodate differing local practices (e.g. different policies regarding the return of findings to subject participants, different standards of security, varying institutional sanctions) and tackle complex consent issues to do with re-use and international transfer of data. The use of cloud-based storage and processing services is becoming increasingly important but it raises issues such as third party access (for example, by security services).
39. The wide availability of information technology and social networking platforms have facilitated participant-led research, allowing individuals to group together to address research questions of interest to them and complement institutional research. The norms and social dynamics of patient-led research are different to more formal institutional research owing to the online medium, self-organising dynamics and the absence of formal review or oversight structures. They present different challenges of ensuring the protection of individual interests, of integration with institutional research, and of translation of findings into clinical products and practice.
40. Good practice is emerging in many population research initiatives but more needs to be done to protect the privacy interests of subject participants in order to secure the trust of current and future generations.

Recommendations

- 10 Appropriate mechanisms should be put in place to allow governance arrangements to evolve during the life of an initiative, through deliberation with morally relevant stakeholders including participants, the public, funders and the research community. Arrangements may include, e.g., representation of relevant stakeholder groups in the governance of the biobank; regular review of a public ethics and governance framework document legitimated through deliberation with interested parties that sets out the relationships of a biobank with participants, the research community, individual researchers, funders and the wider society. This may serve as an instrument to maintain alignment of the public interest in research with the privacy and other interests of the participants. Governance arrangements should, among other matters, outline policies for maintaining data security, the feedback of health-related findings to participants and for research access to the resource. In large scale and complex initiatives detailed diagrams of data flows should be available to support good governance. The responsibility to ensure appropriate governance arrangements are in place rests with funders.
- 11 Where broad consent is sought for the use of data additional, adaptive safeguards should be in place to secure the interests of participants over the life of a project. A possible model is provided by a publicly articulated, 'living' ethics and governance framework that reflects the expectations of participants and is subject to review and revision through mechanisms that involve representatives of the full range of interests of participants in the initiative.
- 12 Researchers should operate demonstrably within a local governance framework able to maintain reasonable surveillance in order to identify inappropriate data use and administer sanctions for misuse. Researchers should be members of a recognised research environment with appropriate arrangements in place to ensure their research meets ethical standards. They should provide undertakings regarding the confidential and secure use of data and that they will refrain from any attempt to

- identify participants from whom data may have been derived.
- 13 All international collaborative data research initiatives should operate within an explicit, public ethics and governance framework that has agreement from the initiative's constituent partners. International collaborators should be able to demonstrate that they can fulfil recommendation 12 by applying equivalently strong governance standards (using legal and other mechanisms available in their national jurisdiction).
 - 14 All partners in international collaborations should integrate the provisions of the ethics and governance framework (EGF) agreed by the initiative as far as possible at their local research site. The partner should ensure that they adhere to the EGF, for example by ensuring participants have given appropriate consent for the use of data and samples in the initiative and that they are informed of potential transfer across borders.
 - 15 National bodies should publish their policies on the use of cloud services in health data settings so that data initiatives can include this in their decision making and interactions with publics and participants.
 - 16 Biomedical researchers should give consideration to arrangements that will maximise the potential of participant-driven research to contribute to generalisable health knowledge and secure public benefits while providing adequate protection of those involved through continuing ethical and scientific appraisal. Key stakeholders are citizen patient researchers, biomedical research bodies, research funders and journal publishers. All stakeholders should encourage optimal use of human studies for improved health outcomes
 - 17 The research community, including all academic and commercial partners, data controllers and custodians, public services and government agencies should actively foster opportunities to create a more explicit and reflective foundation for extending data access in the public interest. We urge all stakeholders in the medical research enterprise to continue to develop robust and comprehensive, yet efficient privacy protecting rules, guidelines and measures. Among other things these should aim at:
 - Providing greater clarity for members of the public about ways that their biomedical data are used, and may be used in the future, along with a realistic acknowledgement that no system can guarantee privacy and confidentiality in all circumstances.
 - Securing commitments from data controllers to a responsible approach to the extension of data access as part of their core mission statement; they must publish information about their approach to data access, transparency and accountability, and whether, and on what terms, they will consider extending access to data.
 - Demonstrable and continual improvement of collection, storage and data access procedures against explicit standards for accuracy, reliability and security

Reflections (chapter 8)

41. Consideration of the state and direction of travel of information technology, data science, research and governance described in the report, and reflections on examples in health care and population research, lead to some practical precepts for professionals involved in data initiatives. In particular they should:
 - identify prospectively the morally relevant values and interests in any data initiative.
 - take special care to identify those people whose interests may be especially at risk, and interests that arise from diverse values

- not rely simply on compliance with the law as a way of securing that data use is morally appropriate, as the law does not always fully reflect moral norms
- identify the existing privacy norms in relation to the contemplated uses of data
- involve a range of those with morally relevant interests in the design of the data initiative in order to arrive at a publicly statable set of expectations about how data will be used
- state explicitly the set of morally reasonable expectations about the use of data in the initiative.
- involve a range of those with morally relevant interests in the continuing governance and review of the data initiative.