ORIGINAL PAPER

Data Sharing and Dual-Use Issues

Louise Bezuidenhout

Received: 20 June 2011/Accepted: 15 July 2011/Published online: 30 July 2011 © Springer Science+Business Media B.V. 2011

Abstract The concept of dual-use encapsulates the potential for well-intentioned, beneficial scientific research to also be misused by a third party for malicious ends. The concept of dual-use challenges scientists to look beyond the immediate outcomes of their research and to develop an awareness of possible future (mis)uses of scientific research. Since 2001 much attention has been paid to the possible need to regulate the dual-use potential of the life sciences. Regulation initiatives fall under two broad categories-those that develop the ethical education of scientists and foster an awareness and responsibility of dual-use issues, and those which assess the regulation of information being generated by current research. Both types of initiatives are premised on a cautious, risk-adverse philosophy which advocates careful examination of all future endpoints of research endeavors. This caution advocated within initiatives such as pre-publication review of journal articles contrasts to the obligation to share underpinning data sharing discussions. As the dual-use debate has yet to make a significant impact on data sharing discussions (and vice versa) it is possible that these two areas of knowledge control may present areas of ethical conflict for scientists, and thus need to be more closely examined. This paper examines the tension between the obligation to share exemplified by data sharing principles and the concerns raised by the risk-cautious culture of the dual-use debates. The paper concludes by reflecting on the issues of responsibility as raised by dual-use as relating to data sharing, such as the chain of custody for shared data.

Keywords Dual-use · Data sharing · Ethics · Responsibility

L. Bezuidenhout (🖂)

Department of Sociology and Philosophy, University of Exeter, St Germans Road, Exeter EX4 4PJ, UK e-mail: Lmb214@exeter.ac.uk

Background

The concept of misuse of scientific information is not new and indeed, hesitations regarding the negative use of research findings have been recorded throughout the centuries. The majority of these concerns, however, involved the direct misuse of information which could be largely solved by controlling access and/or implementing confidentiality agreements.

In the last 100 years advances in science have dramatically increased the scientific research in all areas, which has been accompanied by a heightened awareness of the potential for the misuse/abuse of sensitive information and/or reagents and methodologies. The concept of dual-use came to a fore in the years leading up to the Second World War in the field of atomic sciences (Evans 2010). The awareness that the same piece of research could be used for good as well as bad purposes led many scientists to question the ethical validity of their activities. Thus, the concept of dual-use as the potential for beneficial research to be misused with malicious intent became entrenched in the discussions about atomic research.

Recently, however, shifts in the understanding of technology have widened the formulation of the dual-use concept. Whereas previously technological functions were seen as based solely on the intrinsic physical properties present in a system, it has now been suggested that they should be viewed as a process of evolution involving wider socio-technical systems where different downstream technologies share some of their upstream technological inputs (McLeish and Nightingale 2007). With this shift in focus from implicit to imposed properties comes the awareness that the one piece of research or methodology could potentially contribute to a number of different outcomes in a manner which is dependent on its application in later experiments.

Thus, later formulations of the dual-use concept have grown to include the potential of any piece of beneficial research to contribute towards negative future events, which presents an ethical dilemma for scientists, governments and the security community as they struggle to balance the benefit derived from an avenue of research with the potential to cause harm (Miller and Selgelid 2007). Examples of classic examples of dual-use dilemmas (discussed later in the paper) include the resurrection of the Spanish Flu virus and the publication of its sequence online, and the unintentional generation of a supervirulent mousepox virus through the insertion of the interleukin-4 gene.

Any discussion regarding the possibility of practically addressing the dual-use potential of the life sciences involves a number of complicated issues. These include aspects such as the possible need to regulate beneficial scientific research due to the potential for uncertain, future harm. Furthermore, the fact that dual-use outcomes could involve the convergence of a number of different technologies means that any method to address them cannot be seen as merely putting a stop to one stream of research. In addition, the fact that any risk/benefit analysis is contextual in nature and thus different formulations could exist between interested parties.

It is also necessary that the benefit of any controls be balanced with their potential to affect some of the fundamental norms of science, including freedom of research, openness, and communality (Resnik and Shamoo 2005). Therefore, any controls must confront questions such as the following:

- How can any restrictions be balanced against expectations that often exist for openness?
- How can restrictions be effective while also respecting the fundamental need for freedom in research?
- How are individual and collective responsibilities apportioned for addressing dual-use issues in the life sciences?
- How can any such controls be standardized and regulated internationally?

A 2006 report published by the National Research Council (2006) described the creation of a "culture of awareness and responsibility" which would address the dualuse potential of the life sciences without undermining scientific freedom. Influential in dual-use debates, and contributing to discussions of this "culture of awareness" are two philosophical concepts: the precautionary principle¹ and the doctrine of double effect.² Both these principles assist with addressing the possibility for beneficial research to be potentially misused in the future, as well as the idea that one area of research can embody both beneficial and negative properties. Although considerable debate exists between dual-use scholars as to the appropriate understanding and response to the potential of unintended future harm in the life sciences, these principles contribute towards building a cautious culture of risk aversion which will not hamper scientific freedom. Importantly, both suggest a duty of the scientists to consider the possible future outcomes of their research as well as ways to minimize potential harms.

In addition to improving capacity in ethics education and developing codes of conduct for scientists, there are a number of other initiatives currently being developed to address this aspect of life science research such as developing prepublication review of journal articles, funding reviews, and assessment of personnel and access controls (van Aken et al. 2004; van Aken 2006; Rappert 2010).

One of the most heavily debated of the initiatives is that of pre-publication review which was summarized in a statement released 2003 by the Journal Editors and Authors Group in which they advocated that "Scientists and their journals should consider the appropriate level of design of processes to accomplish effective review of papers that raise such security issues ... [The editors] recognise that on occasion an editor may conclude that the potential harm of publication outweighs the potential societal benefits. Under such circumstances the paper should be modified or not published" (Atlas et al. 2003). This sentiment has been endorsed by numerous other publications, such as the report published in 2004 by the National Research Council entitled *Biotechnology Research in an Age of Terrorism*, which specifically endorsed publication review for what they termed "experiments of

¹ A common formulation of the precautionary principle was proposed at the Wingspread Conference in 1998 and states "where an activity raises threats of harm to the environment or human health precautionary measures should be taken even if some cause and effect relationships are not fully established scientifically".

² The principle of double effect is based in Catholic theology and dates back to Baggini and Fosl (2007). Simply put, it states that actions can have more than one consequence, and not all these consequences are intended.

concern".³ Thus, increasingly, discussions on science are including debates on the limits of freedom and openness in research and the potential for the future misuse of legitimate data.

Dual-use debates have yet to make a significant impact on data sharing discussions, and the rest of the paper considers two issues—firstly whether raw data should be treated in a similar manner to peer-reviewed publications when addressing the dual-use potential of information and how far the obligation to share can be extended in a dual-use context, and secondly whether the responsibilities of the individual scientist and general responsibilities as conceptualized by current dual-use debates are applicable to discussions on responsibility in data sharing.

Discussion

Dual-Use and Data Sharing: Publication and Restrictions

As yet there have been no papers refused publication on the grounds of dual-use potential as part of the open science journal reviews; however the increasing awareness of the dual-use of some aspects of published work is assisting in furthering the cautious culture of responsibility and awareness. It may be that this attitude could differ from that endorsed by those involved in data sharing arguments, where release of data is an obligation. Indeed, in data sharing discussions considerable emphasis has been placed on the need to ensure open access to anyone seeking information. A report on responsibilities of authorship published by the National Research Council (2003) stated that "universal adherence, without exception, to a principle of full disclosure and unrestricted access to data and materials that are central or integral to published findings will promote cooperation and prevent divisiveness in the scientific community, maintain the value and prestige of publication and promote the progress of science".

Thus it appears that possible tensions exist between the dual-use and the data sharing understanding of the obligation to share. This can be examined further with the use of a classic dual-use example.

A Dual-Use Example: Resurrecting the Spanish Flu Virus

In the late 1980s Jeffrey Taubenberger and his team at the Armed Forces Institute of Pathology developed methodology for sequencing genetic material from badly degraded samples. Using this methodology they went on to sequence the entire genome of the influenza strain commonly known as Spanish flu from samples obtained from victims interred in the Alaskan permafrost.⁴

³ The Fink report calls for greater review of experiments that would: (i) Demonstrate how to render a vaccine ineffective. (ii) Confer resistance to antibiotics or antivirals. (iii) Enhance a pathogen's virulence or render a nonpathogen virulent. (iv) Increase a pathogen's transmissibility. (v) Alter a pathogen's host range. (vi) Enable evasion of diagnostic tests. (vii) Enable weaponization of pathogens and toxins.

⁴ This research has been reported in a number of papers including Taubenberger et al. (1997), Reid et al. (1999), and Taubenberger et al. (2005).

This research provided valuable data which would form the basis of further investigations into the extreme virulence of the virus, and marked a significant development in sequencing capabilities. Nevertheless, the submission of their landmark paper for publication caused considerable concern amongst reviewers and editors for the dual-use potential of the research. As the Spanish flu virus was responsible for infecting a third of the world's population, killing an estimated 40 million people between 1918 and 1919 before vanishing from existence, one of the most fundamental questions was whether releasing the sequence of the Spanish flu virus and the details of the methodology would potentially contribute towards the potential misuse of this research by a third party.

While this research was ultimately published, and has since formed the basis of valuable further research into the evolution of the Spanish flu virus and the reason for its extreme virulence it continues to be the subject of considerable debate and remains an example of dual-use research. In particular the release of the Spanish flu genome into open access databases is a cause of much discussion (van Aken 2007).

One of the major areas of debate regarding this paper was whether the danger posed by the dual-use potential of the knowledge would be better minimized through controlling access to the information through publication restrictions on the paper and the sequence, or whether maximizing the number of people aware of the research would facilitate increased surveillance and awareness (van Aken 2007). The Spanish flu paper was subjected to a pre-publication review based on its dual-use potential, and was the first paper to undergo this process initiated by the National Science Advisory Board for Biosecurity (NSABB)⁵ specifically to assess papers which had the potential to pose security risks.

One of the challenges of dual-use discussions is that they directly probe the concepts of freedom and openness in scientific research. While the advancement of science is premised on these norms, much discussion currently exists regarding their limits in scientific research (Rappert 2007). Nevertheless, despite such hesitations, any idea of regulating the dual-use potential of life science research must involve an analysis of the balance between the concepts of freedom and control. If there is research that has the potential to contribute towards future harm (regardless of intent), can it be ethically justifiable to allow this information to be freely accessed and available?

The cautious approach to data release and sharing as advocated by the concept of dual-use could be seen to be at odds with the principles increasingly being advocated by data sharing communities. While both the dual-use and data sharing debates explicitly endorse upholding scientific openness and freedom of research, it may be that differences in understanding of these terms and what they entail may be

⁵ Following the publication of the Fink Report, the Secretary of the Department of Health and Human Services (HHS) created the National Science Advisory Board for Biosecurity and chartered it in March 2004. The NSABB responsibilities include many of the recommendations suggested by the National Academies. The NSABB charter outlines a number of responsibilities including advising on national policies governing publication; public communication, and dissemination of dual-use research methodologies and results; advising on national policies governing local review and approval processes for dual-use biological research, including the development of guidelines for the case-by-case review and approval by Institutional Biosafety Committees (IBCs); and advising on criteria and processes for referral of classes of research or specific experiments by IBCs to the NSABB for guidance Shea (2007).

the source of future ethical tensions. It may be that the obligation to share as promoted by the data sharing community will be strengthened by an obligation towards dual-use awareness.

An additional area of concern which this example raises is whether a differentiation should be made between raw data and peer-reviewed publications. Examples of dual-use dilemmas often use completed research programmes to demonstrate the dual-use potential of the life sciences and neglect to confront the fact that large-scale and high-throughput data analyses are increasingly involving electronic processes that involve sending information to and from scientific databases (Toronto_International_Data_Release_Workshop_Authors 2009) often prior to publication of associated papers. These recent shifts in data sharing principles and practices have led to the rapid accumulation of vast amounts of data often through largely electronic processes, preceding or keeping pace with analyses.

In particular it is thus expedient to question whether the pre-publication release of raw data and the accumulation of such data in databases may have the potential for future dual-use potential. In addition to the problems associated with the limits of informed consent, stigmatization or data misuse, it is possible that these highly beneficial databases may provide information ultimately contributing towards the malicious misuse of science by a third party, and thus to future dual-use dilemmas, raising questions about control.

Thus, comparisons between the risk-cautious culture being advocated by the dual-use community and the strong general approbation of unfettered data sharing (Pryor 2009; Schofield et al. 2009) may raise some interesting areas of discussion. It is feasible to question whether scientists may be increasingly confronting policies of review for dual-use issues while at the same time being explicitly encouraged or mandated to engage in data sharing with as little restrictions on the data as possible. Such seemingly contradictory policies could appear to place different value on research and raw data, and potentially create ethical tension for scientists, who on the one hand are being encouraged to adopt a more risk-conservative attitude to their research, while at the same time being mandated to share data and with the explicit intention of creating opportunities for the un-restricted re-use of the data by other researchers. Thus, further examination into these two cultures and their intersection (in the areas of publication and access to information particularly) will be beneficial for both areas of debate.

Individual Versus Collective Responsibility

Related to the idea of possible restrictions in scientific research are the concept of responsibility and the involvement of the individual scientist. This is discussed through another classic dual-use example.

A Dual-Use Example: The Mousepox Experiments

In 2001 scientists Australian National University were working with a mousepox virus to create a vaccine which would induce infertility in rodents in order to address the problem of rodent-related crop destruction in the agricultural sector

(Jackson et al. 2001). During the course of their research they discovered that the insertion of the interleukin-4 gene into the mousepox virus created an extremely virulent strain of mousepox that could kill even vaccinated mice.

The researchers recognized that there was a remote possibility of this research being adapted for use in human smallpox as a biological weapon. Nevertheless, despite their reservations and with the advice of various advisory boards they published the research (Selgelid and Weir 2010). When published, however, considerable discussion was initiated regarding the publication of this research and the responsibilities of the scientists involved.

In particular a number of articles (such as Nowak 2001) mentioned that the scientists had decided to share their results as a means of warning the scientific community of these unexpected outcomes. The scientists involved recognized a sense of responsibility towards the misuse of the research with Jackson quoted as saying, "We wanted to warn the general population that this potentially dangerous technology is available, We wanted to make it clear to the scientific community that they should be careful, that it is not too difficult to create severe organisms" (Nowak 2001). Nevertheless, the identification of a research topic as containing dual-use potential, the decision to share and the method of sharing are all complicated subjects that hinge on the interpretation of individual and collective responsibility.

A major discussion points in debates on dual-use controls is how to apportion the responsibility for developing, implementing and executing any initiative (Miller and Selgelid 2007). One of the underlying issues in this problem involves the manner in which individual and collective responsibilities are attributed to scientists. As dual-use involves the future potential for research to be misused by another, intent, culpability and moral connections between the researcher and their data all need to be thoroughly examined, and thus the extent of individual and collective responsibilities.

To refer to the dual-use example above, the generation a mousepox virus capable of killing even vaccinated mice carries with it the dual-use potential of future misuse by a third party. It is however unclear as to how the responsibility for the control of this potential and the possible culpability in the event of a nefarious occurrence would be apportioned. While the scientists can be seen as having some level of responsibility towards their research, the entirety of the responsibility for controlling the dual-use potential of their research cannot and should not lie with them.

A recent paper by Kuhlau and colleagues assessed the extent to which individual responsibility can be attributed to life scientists in dual-use dilemmas. Importantly, they suggested that individual scientists should not be required to estimate the probability and magnitude of misuse, as it is beyond the scope of the individual researcher's capacity (Kuhlau et al. 2008).⁶ In highly complex situations such as dual-use dilemmas, individual responsibility is important and should be fostered

⁶ Instead they proposed conditions for responsibility, so that the individual's responsibility is altered by their social role or position as well as their professional capacity and ability to act. Furthermore, the harmful consequences should be reasonably foreseeable, the benefits of the research should outweigh the risks, and that the results should not be reasonably achieved through other means.

through education and awareness however an individual scientist is insufficient for taking responsibility towards controlling dual-use issues.

Thus, using Kuhlau's argument, although responsibility does not wholly rest on the shoulders of the individual scientist, they bear a degree of responsibility towards addressing dual-use issues in the life sciences. Interestingly this awareness of responsibility towards personal work is reflected to some degree in discussions of data sharing. In a study of formal and informal data sharing in different disciplines of scientists researchers in Edinburgh showed that knowledge transfer and communication covered not only the traditional publication of scientific papers and regular dissemination of information about work through presentations, but also raw data from experiments (Pryor 2009). The range and scale of these less formal outputs raised considerable concerns in scientists about who to share with, what to share and why. Thus, scientists were already selecting information to disseminate, suggesting that they felt "a keen sense of "ownership" towards their data, which frequently emerged as an attitude resonant of protectiveness. [The scientists felt] responsible for the data they have generated and are genuinely concerned for the consequences of someone outside their immediate research orbit applying any inappropriate analyses" (Pryor 2009).

Furthermore, in data sharing the collection of raw data is most often associated with a custodial responsibility for the data based on informed consent obtained and commitments to study participants. Discussion exists on how best to preserve this custody once the data is released and accessible for use by a third party. The "chain of custody" and the individual and collective aspects of the responsibility to preserve it are therefore very important for the preservation of the integrity of the individual researcher, the participants and the research process. Thus, in both data sharing and dual-use debates there is a tension between the extent of responsibility of the individual scientist, and the ambiguous responsibility of releasing data for re-use (or use).

Thus, discussions of data sharing already reflect an individual concern regarding downstream (mis)use of data. As dual-use prompts us to rethink how risk is conceptualized—from individual instances of research to cumulative developments in the life sciences (Rappert 2008), this area of intersecting individual responsibility seems poised for including dual-use debates in data sharing discussions and vice versa.

The increasing complexity of science and the division of labor has led to not only narrowing of the areas of individual responsibility, but also a single individual having a number of different roles, and has led to many suggestions of collective responsibility in dual-use debates (von Schomberg 2007). Thus, as a collective the scientific community is increasingly being required to address dual-use issues in research through involvement in debates, policy formation and education with the aim of developing a "culture of awareness and responsibility"⁷ within the life sciences. Extending the dual-use debate to include data sharing issues will create

⁷ Committee on Advances in Technology and the Prevention of Their Application to Next Generation Biowarfare Threats, National Research Council (2006).

further responsibilities for the science community as they consider the dual-use potential of databases and potential means to address these issues.

Summary

The recent increases in data sharing in the life sciences results in a vast amount of raw data being available online for re-use. The benefits of these databases to science are enormous and immediately apparent as these databases repeatedly facilitate novel research. Nonetheless, the principles of data sharing are subjects of considerable debate and attempts to improve the ethical integrity of data sharing are constantly underway.

Value-neutrality in science is rapidly losing ground, and much discussion is beginning to emerge regarding the responsibilities of scientists towards their work (Rotblat 1999; Rappert 2007). The concept of dual-use forces us to confront the idea that beneficial scientific research (and data) has the potential to be misused, and thus extends the conceptualization of risk from individual to cumulative developments. This concept also forces a critical benefit assessment and an acknowledgement that biomedical research cannot, by default, be considered only beneficial (van Aken 2006). Thus, on an individual and collective level, scientists are required to consider this concept in relation to their work.

This understanding of risk in scientific research should be extended to the generation of huge reservoirs of raw data through data sharing which have the potential to be reused for negative as well as positive purposes. Current discussions on dual-use issues are less commonly concerned with raw data than with complete experiments and thus the interface between dual-use issues and data sharing is an area requiring considerable growth on both sides. The ethical tensions potentially existing between the obligation to share and the responsibility of addressing dual-use issues require further examination which would lead to reflection of these issues in educational initiatives, codes of conduct and policy initiatives which are being developed by both parties.

Particularly relating to data sharing, the concept of dual-use will be valuable when reflecting on who should be responsible for the possible dual-use of raw data in data banks, and what can be done to minimize this potential. When developing future dual-use/data sharing controls and educational initiatives it will be important to address these responsibilities in a manner which supports the openness and freedom of research but does not neglect the risk-cautious culture of dual-use or the obligation to share of data sharing. Further debate will be useful to examine how dual-use and data sharing initiatives can be reconciled so as not to place the scientist in an untenable ethical position with regards to their data.

Conflicts of Interests The author declares that she has no conflicts of interests.

References

Atlas, R., Campbell, P., Cozzarelli, N. R., et al. (2003). Statement on scientific publication and security. *Science*, 299(5610), 1149. Baggini, J., & Fosl, P. S. (2007). The ethics toolkit. Oxford: Blackwell Publishing.

- Evans, N. G. (2010). Dual-use bioethics: The nuclear connection, research report for the wellcome trust project on 'building a sustainable capacity in dual-use bioethics'. Bradford: University of Bradford.
- Jackson, R. J., Ramsay, A. J., Christensen, C. D., Beaton, S., Hall, D. F., & Ramshaw, I. A. (2001). Expression of mouse interleukin-4 by a recombinant ectromelia virus suppresses cytolytic lymphocyte responses and overcomes genetic resistance to mousepox. *Journal of Virology*, 75(3), 1205–1210.
- Kuhlau, F., Eriksson, S., Evers, K., & Hoglund, A. T. (2008). Taking due care: Moral obligations in dualuse research. *Bioethics*, 22, 477–487.
- McLeish, C., & Nightingale, P. (2007). Biosecurity, bioterrorism and the governance of science: The increasing convergence of science and security policy. *Research Policy*, 36, 1635–1653.
- Miller, S., & Selgelid, M. J. (2007). Ethical and philosophical considerations of the dual-use dilemma in the biological sciences. *Science and Engineering Ethics*, 13, 523–580.
- National Research Council. (2003). Sharing publication-related data and materials: Responsibilities of authorship in the life sciences. C. o. R. o. A. i. t. B. Sciences.
- National Research Council. (2006). Globalization, biosecurity and the future of the life sciences. Washington D. C.
- Nowak, R. (2001). Disaster in the making. New Scientist, 2273(13 January), 4-5.
- Pryor, G. (2009). Multi-scale data sharing in the life sciences: Some lessons for policy makers. International Journal of Digital Curation, 4(3), 71–82.
- Rappert, B. (2007). Biotechnology, security and the search for limits. New York: Palgrave Macmillan.
- Rappert, B. (2008). The benefits, risks, and threats of biotechnology. *Science and Public Policy*, 35(1), 37–43.
- Rappert, B. (Ed.). (2010). Education and ethics in the life sciences. Canberra: Australian National University Press.
- Reid, A. H., Fanning, T. G., Hultin, J. V., & Taubenberger, J. K. (1999). Origin and evolution of the 1918 'Spanish' influenza hemagglutinin gene. PNAS, 96, 1651–1656.
- Resnik, D. B., & Shamoo, A. E. (2005). Bioterrorism and the responsible conduct of biomedical research. Drug Development Research, 63, 121–133.
- Rotblat, J. (1999). A Hippocratic oath for scientists. Science, 286(5444), 1475.
- Schofield, P. N., Bubela, T., Weaver, T., et al. (2009). Post-publication sharing of data and tools. *Nature*, 461(10 September), 171–173.
- Selgelid, M. J., & Weir, L. (2010). The mousepox experience: An interview with Ronald Jackson and Ian Ramshaw on dual-use research. *EMBO Reports*, 11(1), 18–24.
- Shea, D. A. (2007). Oversight of dual-use biological research: The national science advisory board for biosecurity. CRS report for congress. C. R. Service.
- Taubenberger, J. K., Reid, Ann. H., Krafft, A. E., Bijwaard, K. E., & Fanning, T. G. (1997). Initial genetic characterization of the 1918 'Spanish' influenza virus. *Science*, 275(5307), 1793–1796.
- Taubenberger, J. K., Reid, A. H., Lourens, R. M., Wang, R., Jin, G., & Fanning, T. G. (2005). Characterization of the 1918 influenza virus polymerase genes. *Nature*, 437, 889–893.
- Toronto_International_Data_Release_Workshop_Authors. (2009). Prepublication data sharing. Nature 461(10 September), 168–170.
- van Aken, J. (2006). When risk outweighs benefit: Dual-use research needs a scientifically sound riskbenefit analysis and legally binding biosecurity measures. *EMBO reports*, 7(Special issue), S10– S13.
- van Aken, J. (2007). Is it wise to resurrect a deadly virus? Heredity, 98, 1-2.
- van Aken, J., Johannsen, S., & Kollek, R. (2004). Die Terror-Angst und ihre Folgen. Deutsches Aerzteblatt, 101(45), A3014–A3916.
- von Schomberg, R. (2007). From the ethics of technology towards an ethics of knowledge policy and knowledge assessment. A working document from the European Commissions Services. European Commissions Services.