
Whose data?

TRANSCRIPT

Research Ethics Online Course

1. This lecture looks into ethical challenges around data ownership.
2. Data is in the heart of research work, in vast majority of cases without data there is no analysis, no conclusions and no publications. As the research work is about analysis, conclusions and definitely about publishing your work, data ownership becomes an important aspect. The ethical challenges around data ownership are typically most clearly recognised and most complicated when it is time to publish results.
3. So data ownership is closely linked to our ethical considerations around who is allowed to publish results based on any given data or data set or in other words claim those results as part of their own degree, academic career, expertise and kudos.
4. There is no direct link between collecting data and having a right of ownership of that data in a sense of having an exclusive right to analyse or publish from the data. Data may be collected by technicians or undergraduate students, for example. It would not make academic sense to state that the right to analyse and publish from the data would in these situations belong to those who have collected it. Alternatively data an individual researcher or a research team collects is more and more often part of a larger collective data set, where boundaries of ownership are almost by definition no longer related to the actual collection of data.
5. Further, the research community is increasingly engaging with concepts of open data sharing, which is also set as a high level research principle by the European Union for example in Horizon 2020 program has stated open data sharing as one of the key principles attached to funding. Sharing data before the collectors have been able to publish extensively from it, presents further challenges to data ownership.
6. So how do we form sense of ownership around data particularly in relation to a right to publish from the data? There are no hard and fast rules around data ownership. For this reason, planning around data ownership as early as possible will reduce confusion and ethical challenges later.
7. Let's first look at data ownership issues on a local level, in other words how to decide on data ownership within a research group. Ideally the planning should take into consideration at least the following aspects or if a problem has already developed, the solution should consider the following:

-
8. Independence of the research work - the more independent the work is, the more right to ownership is built. This is not only independence of actual data collection, but more importantly independence in relation to conceptualising the data collecting process in relation to research aims and designing the data collection practice.
 9. Relationship with other projects and data - the more independent the data is from data collected by others in the research group, the stronger the right of ownership is. If the data collected stands as a highly independent and clearly defined data set with only loose and consequential ties to other data collected, the more weight there is for the right to analyse and publish from the data to belong to the one who has collected it.
 10. Contractual relationship with the university / research group - the contractual relationship provide one legal framework for defining data ownership. If the data collector is employed by the university, the data ownership is different from a self-funded student enrolled at the University. It is important to understand data ownership set up by the contract you sign with the University or any other research establishment so that you will know your rights and avoid any surprises later.
 11. Copyright to own work (analysis, words, images) - copyright law provides another legal framework to look at data ownership. Typically copyright is applied to analysis and the subsequent words or images from the data, rather than the data itself. So while you have a right to the products from your data, copyright does not give you similar legal ownership of the data itself.
 12. Fair expectations by the employer on workload - your right to publish from your data cannot be taken away from you if you have not been given a fair opportunity to publish from the data yourself. The employer needs to consider fairness in the manner it allocates workload to different members of the research group. So if you feel that your data is taken away from you because you have not published from it quickly enough, while the you have been put under pressure to perform other duties within the group, it could be a sign of an unfair expectation on workload and it should not be used to reduce your rights to publish from your data. At the same time, it is important to note that members of the group have responsibilities towards the group and there is a balance to be drawn between these responsibilities and rights of ownership.
 13. It cannot be emphasised enough to discuss these aspects within the research team to avoid challenges. As the group develops and individuals progress through their research plans with all the unavoidable changes, the data ownership plan should be reviewed regularly by the team.
 14. Data management agreements including data storage and usage plans made at the start of the project will influence the interpretation of data ownership. So if you do not know what these agreements are, start a conversation with your supervisor and colleagues to uncover what is assumed in the group and develop a better understanding and maybe even more formal agreement on how the group will

manage these issues in the future. Below are case studies of data ownership problems, which have surfaced when no data ownership plans were in place.

15. Now let's move onto considering data sharing from a more global perspective. The movement towards open sharing of research data has its roots in the 1940s, while it has truly gained momentum in after 2000 with multiple agreements, statements and principles now in place to increase data sharing within the research community and with the society in general. Stop and think, what do you think is behind the open data sharing movement, why do you think it is important (or not)? Pause the lecture, write your thinking down or make a mental note and then press play again to see what others have thought.
16. Here are reasons given in the EU Horizon 2020 project, which echo many of the other initiatives closely:
 - i. build on previous research results (improved quality of results);
 - ii. foster collaboration and avoid duplication of effort (greater efficiency);
 - iii. accelerate innovation (faster to market = faster growth);
 - iv. involve citizens and society (improved transparency of the scientific process).
17. Ethically the open data sharing makes clear sense when looked at consequentialist perspective and considering all the stakeholders or the community as one unified stakeholder. The benefits of open data sharing for the collective of stakeholders are significant in comparison with the alternative of no data sharing and the harms appear to be minimal as long as issues like privacy and security are appropriately considered.
18. The ethical challenges appear more clearly though when we consider this approach from a more individual perspective. Imagine yourself in the data collection phase of your project - it is often laborious and slow, if not even quite monotonous. Now move on to thinking that all your data is in, you have got it in a format you can start analysis from which to draw conclusions. Your data is rich and you have plans to publish at least three articles from it. Now imagine sharing all your data with the first article. How would you feel? Scared? Proud? Worried? These feelings would be based on the fact that you are fully aware that in the academic world the main measure of ability, effort and career progression is your publications. Thus, the benefit of collecting data without the associated ability to publish from it presents a potential for an unfair system where some are able to gain career advantage from the labour of others while those who have collected the data are left with no academic merit for having done so. So we have a conflict between the benefit of many compared to the fairness towards a few. As the inability to publish from your data would change the way people approach and engage in data collection, the academic world seeks to find ways to support and maximise the social advantage of open data sharing, while maintaining a fair division of academic merit. These include timing the release of data and developing concepts like stewardship of data that allows priority for publishing from the data.

19. There is no doubt that open data sharing together with open access publications (which are discussed in more detail in relation to sharing research data later in this course) are here to stay and they will change the way we view our research practice and the way we collaborate in a profound way. Taking this into consideration from the start of your research project will put you in a good position to both to grow your academic career and to make sure your work has the maximum potential to benefit the society.