Finding the Law for Sharing Data in Academia

Esther HOORN[a,1] and Marlon DOMINGUS[b,2]

[a] *Rijksuniversiteit Groningen, The Netherlands*
[b] *Erasmus University Rotterdam, The Netherlands*

Abstract. How can universities provide good advice about the legal aspects of research data management? At the same time, how can universities prevent that perceived legal risks become barriers to: conducting research, sharing research data, valorisation of research data, and control mechanisms for the purpose of scientific integrity? A Dutch expert group developed a creative approach based on some core ideas[3] about regulation in the field of academic research.

Keywords. Research, hard law, soft law, code of conduct, guidelines, model contracts, wiki

1. From Self-funded Science to Publicly Funded Academia

History shows that the funding of research has an impact on research itself. From a Renaissance culture with its roots in self-funded science, via noble and religious patronage, to government funding, military funding, patent profits, corporate sponsorship, and private philanthropists, researchers have found ways to contribute to science or scholarship as well as meet the requirements of their research grant providers.

Today, Academia[4] in the Netherlands is mainly publicly funded[5] but in recent years the so called *third flow of funds* has seen a substantial increase.[6] This refers to revenues based on contract research and funds from Dutch ministries and the European Union (FP7 and Horizon 2020).

With the third flow of funds comes the obligation to meet the new funder's requirements. The EU grants, for instance, are aimed at fostering interdisciplinary re-

---

[1] Corresponding Author. E-mail: e.hoorn@rug.nl
[2] Corresponding Author. E-mail: domingus@ubib.eur.nl
[3] The views expressed here are those of the authors and do not necessarily reflect those of the Rijksuniversiteit Groningen or Erasmus University Rotterdam.
[4] Used as an equivalent to the 14 research universities in the Netherlands. The case for private universities and university colleges is not investigated here.
[5] Directly and indirectly funded by the government. Direct funding from the Ministry of Education, Culture and Science, and Wageningen University's funding comes from the Ministry of Economic Affairs (the so called *first flow of funds*). Indirectly funded by the government: grants from the Dutch Organisation for Scientific Research (NWO) and the Royal Netherlands Academy of Arts and Sciences (KNAW) (the *second flow of funds*). See also: http://www.vsnu.nl/funding-en.html and http://www.rathenau.nl/fileadmin/user_upload/rathenau/De_Nederlandse_Wetenschap_Facts_and_Figures-Dutch_universities_2012__01.pdf
[6] From 22% (2004) to 26% (2013) of the gross income of Dutch research universities. Online source: http://www.vsnu.nl/funding-en.html

search, international collaborations as well as Industry-Academia Partnerships. And the EU runs a pilot on open data:

"*The European Commission's vision is that information already paid for by the public purse should not be paid for again each time it is accessed or used, and that it should benefit European companies and citizens to the full.*" [1]

So Academia's move to the public domain for research funding has resulted in a new Academia ecosystem in which new (consortium) agreements and understandings are to be formulated.
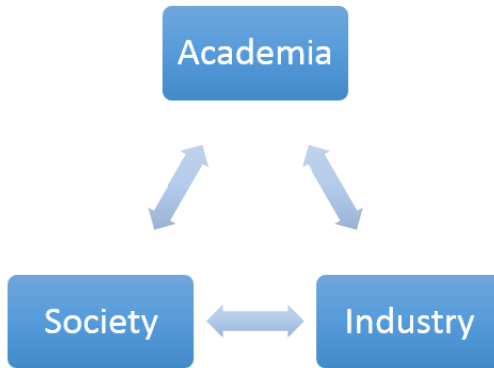

Figure 1. Academia Ecosystem.

Signs of the new ecosystem are, for instance, Academia's focus on 'societal impact' and 'societal relevance' [2]. In the new ecosystem, the moral pressure to make publicly available what was publicly funded, relates to both research publications and the underlying research data. Dissemination of research data is demanded from the researcher rather than exploitation of the data.

The parallel with a company's shareholders is clear: government invests in Academia with citizens as the Academia shareholders, advocating citizen science. Academia's capital is arguably the collective research data, which combined with government's open data and the industry's data open all kinds of new possibilities for all in the ecosystem.
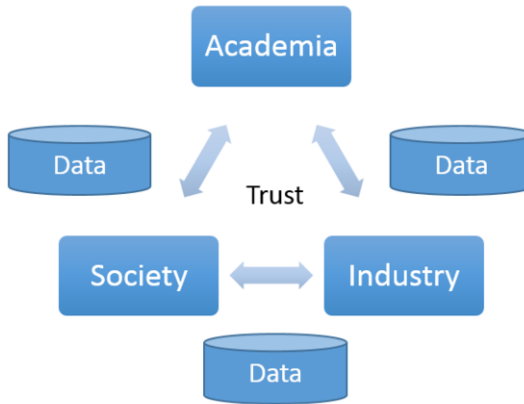
Figure 2. Data sharing in the ecosystem.

In this new ecosystem new rules and agreements are required, new responsibilities assigned for the ecosystem to find its balance. Data exchange seems only possible based on mutual trust. How can universities best provide good advice to their researchers about the new legal and ethical aspects?

2. Hard Law, Soft Law and Ethics

Whereas in some fields of law norms and procedures are crystal clear, in other fields of law open norms leave bargaining power to participants. When considering a good approach for raising awareness on legal aspects of sharing research data, this fuzzy approach to law seems to match the nature of law in the field of academic research. In addition to the criteria of research funders, a broad range of hard law, like privacy regulation and contract law and soft law (opinion juris), like research codes and discipline-specific norms are applicable in decisions about openness and involvement of citizen in science. A code of conduct is not a body of law, but a canon for self-regulation, based upon ethical principles [3]. For Academia these principles are summarized as: Responsible Research and Innovation, for Society: Good Citizenship and for Industry: Corporate Social Responsibility. When research is considered, we distinguish different ethical dimensions [3]: with regards to the *context of research*[7] and the *responsible conduct*

---

[7] ALLEA, p. 10: "Could the research result in harm for people, nature or society, or be in conflict with basic human values?" This aspect, however, is ignored in the ALLEA Code. R. von Schomberg, European Commission-DG Research and Innovation, introduces the following normative anchor points for this ethical dimension: 1. Compliant with fundamental rights. 2. Sustainable and 3. Socially desirable. See: Von Schomberg, Prospects for Technology Assessment in a framework of responsible research and innovation, M. Dusseldorp and R. Beecroft (eds). *Technikfolgen abschätzen lehren: Bildungspotenziale transdisziplinärer Methoden* (2011),Wiesbaden: Vs Verlag.

of research.[8] In conclusion: good advice to researchers addresses hard law as well as soft law and takes into account the two ethical[9] dimensions.
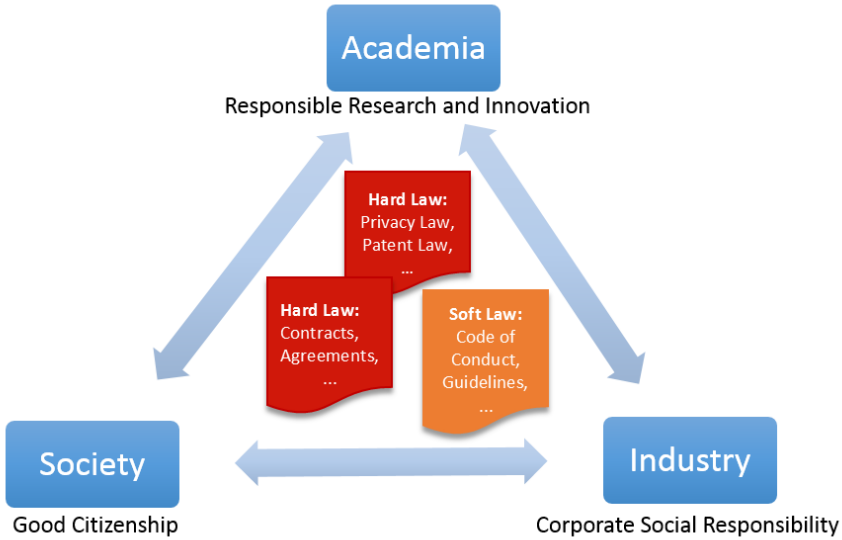

Figure 3. Hard law and soft law in the ecosystem.

At the same time, an approach that depends on specific national or local regulations (hard law) is not considered fruitful, especially since research practices tend to transcend borders, borders between different legal systems with respect to legal aspects of data management. For instance, in the Netherlands ownership of data is not well defined in law. This might be a blessing in disguise to stimulate discussions about data stewardship [4] and shared responsibilities in data management practices. For legal practitioners, however, new solutions are sought within the system of law starting from the relevant and vigorous field of IP law in which the institution is the rights holder.

It's interesting to observe here that principles of scientific integrity are perceived to have a universal character, whereas the different formal legal systems as well as the different good practice rules have national boundaries.[10]

---

[8] ALLEA, p. 10. This amounts to definitions of proper scientific practice and of scientific misconduct, based upon principles of scientific integrity, and guidelines for good practice rules.

[9] Von Schomberg suggests to regard ethics as a "design" factor of technology and increase social-ethical reflexivity in research practices by incorporating ethical principles in the design process of technology (privacy by design as an example), which can lead to well accepted technological advances. See p. 15: Von Schomberg, Prospects for Technology Assessment in a framework of responsible research and innovation, M. Dusseldorp and R. Beecroft (eds), *Technikfolgen abschätzen lehren: Bildungspotenziale transdisziplinärer Methoden* (2011),Wiesbaden: Vs Verlag,.

[10] See ALLEA: *The European Code of Conduct for Research Integrity*, p. 9.

3. Dimensions in Data Openness

Data openness is yet another dimension which needs to be addressed here. In our eco-system we would expect there to be data sharing between Academia, Society and Industry, based on a mutual trust, once agreement has been reached on relevant aspects of hard law and soft law. This data openness should be regulated and well defined, but open. We see, however, different dimensions in data openness.

The European Commission supports open data. Open data refers to the idea that certain data should be freely available for use and re-use [5]. More precisely, open data is the engine for innovation, growth and transparent governance [6]. Furthermore, open data is supported by the European Commission in the context of open science[11] and citizen science[12].

Similarly the Netherlands Organization for Scientific Research (NWO) holds [7] that research results paid for by public funds should be freely accessible worldwide. This applies to both scientific publications and other forms of scientific output. In principle, it should be possible to share the research data with others as well. In this way, valuable knowledge can be utilised by researchers, businesses and civil society organisations.

From the point of view of Academia[13] however, openness seems more restricted to fellow researchers (interested colleagues) from whom the general public benefits through their publication of research findings, thus contributing to public knowledge. For this reason, research data should be available to colleagues who want to replicate the study or elaborate on its findings.[14] To be realistic: within Academia the role of the individual researcher and his / her motives for not creating open data should not be underestimated. Within many disciplines, sharing too much data too soon[15], could endanger an academic career.

---

[11] See the EC policy on Open Science: The European Commission has promoted an approach to research and innovation in which all societal actors (researchers, citizens, policy makers, businesses, civil society organisations, etc.) work together during the whole Research and Innovation process, with the aim to better align research and innovation outcomes with societal values needs and aspirations. It has referred to this approach as Responsible Research and Innovation (RRI). Source: http://ec.europa.eu/research/swafs/index.cfm?pg=policy&lib=science

[12] See the definition in the 2014 EC Green Paper on Citizen Science [*Citizen Science for Europe. Towards a better society of empowered citizens and enhanced research*]: "Citizen Science refers to the general public engagement in scientific research activities when citizens actively contribute to science either with their intellectual effort or surrounding knowledge or with their tools and resources. Participants provide experimental data and facilities for researchers, raise new questions and co-create a new scientific culture. While adding value, volunteers acquire new learning and skills, and deeper understanding of the scientific work in an appealing way. As a result of this open, networked and trans-disciplinary scenario, science-society-policy interactions are improved leading to a more democratic research based on evidence-informed decision making." Online source: http://ec.europa.eu/information_society/newsroom/cf/dae/document.cfm?doc_id=4121

[13] See ALLEA (p. 10, §2.2.3.) the: *Open Communication* principle.

[14] See ALLEA (p. 13, §2.3.) *Good data practices: availability and access*.

[15] Sometimes it can *never* be the case that research data comes available; to protect the privacy of patients and / or to protect the commercial interests of Industry. This is the case with most medical research, for example in the field of epidemiology.

4. A Data Case

In the case of the recent NWO data contract (April 2015) NWO states that researchers are in some cases (depending on field of science and type of the awarded NWO grant) obliged to enter into a data contract with Data and Archiving Services (DANS), the NWO service provider for research data archiving. The data contract is intended to guarantee accessibility to the data as well as digital sustainability of the data for additional scientific research [8]. As we have learned from NWO's open data position, however, the data should in principle be available not only to researchers, but also to businesses and civil society organisations.

At the same time the granting conditions of NWO recognise the research funder's and the research institution's shared ownership of the research data. It is evident that the data management of a project must also be in line with the institutional policies and responsibilities regarding research data.

Moreover, the interests of the partners involved in the project also need to be taken into account. In this respect, ethical commissions play an important role by shaping self-regulation for research integrity via peer assessments. The institution has a legal obligation [9] to ensure that a proper and independent assessment framework is in place to assure recommendations are followed up. Finally, when personal data is involved, the institution is responsible for technical and organisational measures (privacy by design) to ensure privacy of participants, also during and after the project.

Analysing the contract from a lawyer's perspective the contract seems to imply that the project-leader can bind the institution to sign a license after the project. This is not the case. It would be good practice to make this explicit in the contract.

5. Approach: Legal Research Support

So, the research data landscape is altogether a complex one. In any given case, ethical, legal and social implications can be identified, as we have seen. These implications may be perceived as / or may actually be barriers to:

- conducting research,
- sharing research data,
- valorisation of research data and
- control mechanisms for the purpose of scientific integrity.

How can we lift these barriers? In essence: all those concerned need to have a suitable understanding of the matters. Some, however, should acquire expert's[16] knowledge on these issues and should act as the go-to person for identifying what the relevant aspects are, what the relevant ruling is, what the course of action should be and what agreements, contracts or otherwise need to be formulated.

Currently in Academia in the Netherlands, research support services is a joint effort of staff from Faculty, ICT, Library, Legal Affairs, Academic Affairs, Valorisation Offices and Patent Offices in many different roles. As is the case with most hot topics, there are many perspectives, opinions and interests. This is immediately clear once you

---

[16] Rob Posthumus suggested the term 'consciously incompetent' in this respect, derived from the psychological "conscious competence" learning model, usually attributed to Abraham Maslow. It is the expert's task to be 'consciously competent'.

engage in a discussion on the topic of research data management. But it is also a topic no single person is likely to get their head around as it requires co-operation.

We propose not so much a staff / responsibilities matrix, but rather suggest that within a university the following basic steps should be taken by experts:

1. Identify barriers and pitfalls, for instance in a research project plan,
2. Acquire accurate knowledge of the rules and requirements regarding to research data,
3. Take proper legal advice and applying this legal knowledge correctly and timely, for instance by writing tailored paragraphs in a Consortium Agreement.

Even within a single university this is a challenging task. The co-operation between universities to collectively build a body of knowledge, best practices and model agreements looks promising. From a researcher's perspective, it should be clear whom to turn to for support related to these matters. We suggest implementing an awareness program in which the researcher is offered an overview of the research support services and choose[17] from them when specialized support is needed.

From the expert point of view, one would expect an ongoing process of creating a structured body of knowledge, resulting in a detailed analysis of the legal and ethical requirements and the corresponding best practices, template paragraphs and model agreements.[18] On a research data management services level, this list could be considered as a basic functional roadmap to ensure that what is agreed within a research project, is actually executed as promised.

6. The Wiki and the Mood Board

In a seminar [10] addressing these matters, a Dutch expert group presented a wiki [11], in which the three basic steps as described above form *la ligne rouge*.

As a tool to further discussions about legal aspects of data management, the diversity of perspectives and approaches to regulation is visualised on a map. This Mood Board [12] is also a playful way to identify domain specific legal and ethical barriers and pitfalls.

The group is now setting up a broader network with legal practitioners of the institutions and hopes to set, with you, an agenda for the development of helpful support material.

---

[17] And in this respect following *The Netherlands Code of Conduct for Academic Practice. Principles of good academic teaching and research*. Association of Universities in the Netherlands (VSNU), 2014. Online source:
http://www.vsnu.nl/files/documenten/Domeinen/Onderzoek/The%20Netherlands%20Code%20of%20Conduct%20for%20Academic%20Practice%202004%20(version%202014).pdf

[18] The wiki mentioned in the next paragraph is, in our view, a candidate of a platform of such a body of knowledge. See for instance the section: *The landscape of present rules and requirements regarding to research data, for instance in the recently revised Code of Conduct for Scientific Practice and in the regulations applied by research funding bodie*s: https://wiki.surfnet.nl/pages/viewpage.action?pageId=47449662

Figure 4. Mood Board.

Acknowledgements

References

[1] Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020, Version 1.0. 11 December 2013. Online source: http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf.

[2] R. Owen, Responsible research and innovation: From science in society to science for society, with society, *Science and Public Policy* **39**(6) (2012), 751–760.

[3] All European Academies (ALLEA), The European Code of Conduct for Research Integrity, 2011. p. 8 Online source: http://www.esf.org/fileadmin/Public_documents/Publications/Code_Conduct_ResearchIntegrity.pdf

[4] D. Kleppner, Ensuring the integrity, accessibility, and stewardship of research data in the digital age, *International Association of Scientific and Technological University Libraries, 31st Annual Conference*, (2010). Online Source: http://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=1023&context=iatul2010

[5] http://ec.europa.eu/digital-agenda/public-sector-information-raw-data-new-services-and-products

[6] European Commission, Open data. An engine for innovation, growth and transparent governance, 2011. http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0882:FIN:EN:PDF

[7] Online source: http://www.nwo.nl/en/policies/open+science

[8] Online source: http://www.dans.knaw.nl/en/deposit/information-about-depositing-data/datacontract/data-contract?set_language=en

[9] Online source: Artikel 1.7, Wet op het hoger onderwijs en wetenschappelijk onderzoek. Online source (Dutch only): http://wetten.overheid.nl/BWBR0005682/Hoofdstuk1/Titel1a/Artikel17a/

[10] A. van de Wijngaart, *Mapping ownership in the data landscape*. Online source: https://wiki.surfnet.nl/download/attachments/47449647/Seminar%20Report%20RDO.pdf

[11] Online source: https://wiki.surfnet.nl/display/RD/WIKI+Research+Data+Ownership

[12] Online source: https://wiki.surfnet.nl/pages/viewpage.action?pageId=47449973