Knowledge repositories. In digital knowledge we trust

Tsjalling Swierstra^{1,2} · Sophia Efstathiou²

Published online: 17 September 2020
© Springer Nature B.V. 2020

It is self-evident that the practice of healthcare makes extensive use of scientific knowledge, while also contributing to it. However, the scientific work processes and the people that generate scientific knowledge often remain somewhat in the background of healthcare delivery. It is therefore less visible that when biomedical scientific practice changes, particularly in its methods, devices, and stakeholders, those changes also deeply affect how healthcare is organized and performed.

This special section explores one such pervasive shift in biomedical science-making, and its effects on healthcare stakeholders: the rapidly increasing use of computational tools and resulting data. Somewhat unusually, the special section is spread over two issues: the contributions by Nydal et al. and by Gabrielsen have already been published in issue 23(3), the remaining three contributions constituting the special section you find in the current issue 23(4), right after this introduction.

Though the idea of big data biology arguably dates back to post-WWII Big Science initiatives in natural science and ecology (Aronova et al 2010), the introduction of computers into the experimental and observational practices of biology crucially enabled the genomics revolution. Current bioscience is increasingly digital and—in the terms of philosopher of science Sabina Leonelli—‘data-centric’: it aims to produce and handle data in new collective ways, while preserving enough detail to make data useful in particular and individual contexts (2016). Computational tools have raised important questions about the methods of ‘big data biomedicine’: is bioscience transforming into an engineering discipline “changing the living world without trying to understand it”, as microbiologist Carl Woese asks (2004,

173)? Is this work introducing new ways of thinking and doing science, e.g. by considering knowledge as a computable thing (Efstathiou et al 2019), or is this type of data-centric work a scaling up of past attempts at (perspectival and limited) scientific understanding (Callebaut 2012)? And what are the aims of this work? Are the extensive population data collected as part of ‘personalising’ medicine no more than “promissory data” to be used to achieve good results in some unidentified future (Hoeyer 2019)? Or can we imagine a future for personalised medicine that considers questions of justice and access alongside improved scientific understanding (Prainsack 2017)? Expectations for ‘decoding’ the book of life have been undercut by more messy realities of -omic complexities, but understanding and addressing health and illness on the molecular level is still a live goal for twenty-first century biomedicine, and it is impacting stakeholders.

This special section focuses on a particular technological innovation as a fixed point of relational analysis. It explores how digitalisation in scientific, knowledge-producing practices affects stakeholders of medicine and healthcare. The articles collected here explore how digitalisation influences what skills and virtues are required from the stakeholders who together contribute to delivering healthcare: from professionals like biomedical scientists, lab analysts, data scientists and curators, bioinformaticians, and doctors; but also from lay audiences, like the citizens who are asked to donate their biomaterials and the patients who visit the doctor for help, information and guidance. What we may conclude from these studies is that changes in scientific practice impact the ethos and relations between all these stakeholders, what they expect from one another, how they collaborate and communicate with and depend on one another.

At the heart of this digital transformation lie the so-called knowledge repositories, or bio-banks. Nowadays, biomedical data are produced through high-throughput analytical tools in such a rapid pace and on such a vast scale, that these can no longer be communicated through the age-old, well-established channel of the academic paper. No one can keep up with the latest developments. Also, storing scientific

✉ Tsjalling Swierstra
t.swierstra@maastrichtuniversity.nl

¹ Department of Philosophy, Maastricht University, Maastricht, The Netherlands

² Department of Philosophy, Norwegian University of Science and Technology, Trondheim, Norway

information on paper makes it hard to combine all the findings, and it is exactly in making such combinations that the promise of the data revolution lies. Many people nowadays conclude that data-centric science marks the point where data have to be ‘freed’ from their paper confinement and stored in the cloud. Only there can the larger, overarching, picture become visible. Data also have to be freed from the confines of the individual scientist, whose brain is no longer able to process all the data. Ideally, one might infer, all publications should therefore be computer readable, so that the computer can detect the patterns that escape the human mind.

Indeed, a new paradigm for storing biomedical data is rapidly gaining ground: the so-called FAIR guidelines. (Wilkinson et al. 2016) Data should be Findable (this means that data have clear and uniform labels); Accessible (so not stored in some stuffy journal or in a private archive, but in an open-access digital database that can be consulted from any computer with Internet); Interoperable (it should be uniform, standardised, so they can be endlessly combined into a bigger picture); and data should be Reusable (other scientists should have access to the everyone else’s data for their own usage). In the end, or so goes the guiding dream behind this type of research, this way of collecting and combining and analysing vast amounts of data will lead to an understanding of the billions of molecular pathways that make up a living organism. And in the wake of such understanding, it will become possible to influence, or even to control and manipulate, life from the molecular level up.

This molecular medical imaginary holds great promises, among which those of personalized or precision medicine are the best-known (Hedgecoe 2004, see also De Grandis and Halgunset 2016). In this special section we opt for a different perspective: we look at all the nitty gritty work that has to be performed to make this epistemic shift happen in practice, and at the people and relations affected in the process. This requires much more than more and faster computers, or other forms of technology. It requires a pervasive reconfiguration of the whole field of health care, starting with the domain of biomedical research itself, but eventually rippling through to all other fields and affecting all stakeholders. Scientific knowledge, after all, is not a privilege of the isolated individual, but is produced by a community and is in that sense a result of social relations. (Hardwig 1991, 697) The shift requires a reorganization of relations among the stakeholders who have to be willing to adopt new roles and make new kinds of contributions to the common endeavour. And as with all relations, these relations too have an ethical dimension: what are legitimate expectations among those stakeholders? And what (new) virtues should they learn and what (new) vices should they unlearn?

That this ethical dimension is real, is manifest from how different parties respond to the changed requirements with

which they find themselves confronted. Not everyone is willing to embrace data intensive biomedical science, nor to adapt to the newly prescribed roles. Scientists may for example protest against the reduced role of journal publications—if only because they are subject to a career system that is still largely built on having such publications in your name. Or they may object that they don’t have the time or skill to make their findings computer-readable. Or they may be unwilling to share all their hard-won data freely with anonymous others. Or they may feel that the ‘one size fits all’ approach of standardized knowledge repositories actually hampers their own scientific efforts. Other stakeholders who have protested against the new roles and responsibilities, are the citizen-donors who are requested to donate their data – in the form of genetic material – for the common good. They have questioned whether—and if so, under what conditions—there indeed exists an ethical obligation to donate such sensitive materials, and whether the costs and benefits of data science are distributed justly. Doctors grapple with the new data-intensive science when they relate to their patients: how best to communicate this type of finding, and how to combine one’s duty to care with a duty to respect the autonomy of the patient? And not all patients are enthusiastic about the new tasks and responsibilities the new approach to their health implies. All such conflicts, struggles and negotiations are an indicator that in the case of data-intensive science we are not merely dealing with epistemic questions regarding true scientific knowledge, or with technological questions about how to organize the production of digital knowledge, but also with ethical and political ones.

Many of such ethical issues are touched upon in this special section. However, the shared focus of all five contributions lies on ‘trust’ – as a relational challenge, as an organizational challenge, and as a personal disposition or even virtue. Trust can be seen as a kind of meta-value, in the sense that it provides an environment in which ethical values and virtues tend to flourish. In the words of Sissela Bok: “Whatever matters to human beings, trust is the atmosphere in which it thrives. (Bok 1978, 31) Trust has been widely discussed by both philosophers and social scientists. Among them, two authors loom large in the following contributions. The first is a philosopher, Annette Baier (1986). From her we derive the crucial insight that trust is not—as writers before her had argued—a contractual relationship between equals, but indicates a relation of asymmetry and is inseparable from vulnerability. Trusting someone implies that one is aware that by doing so one runs the risk of being hurt by the other. In doing so, one acknowledges an asymmetry of power and control. Trust means that we muster the courage to willingly depend on the good-will of others. The second author is a sociologist, Niklas Luhmann (2000). From him we derive the distinction between confidence and trust. Confidence is a form of willingly depending on others, but as a

matter of routine. A defining feature of confidence is that it is pre-reflexive and doesn't require reasons or justifications. We don't trust traffic lights to function appropriately, but we are confident that they do so. We don't trust that we walk the corridors of our laboratories without risking physical violence, but we are confident that any violence we may encounter will be psychological rather than physical. By contrast to confidence, trust is more active and deliberate, and points to the willingness to take a risk. Trust is putting your fate in the hands of someone else, in the mechanics of something else, when you are well aware that this may backfire. We do reason about whether or not our trust in someone or something is indeed justified. Luhmann sums up the difference between confidence and trust nicely: when something goes wrong, in the case of 'confidence' we typically blame the other; in the case of trust we tend to blame ourselves ('I shouldn't have put my trust in....'). In sum, whereas confidence marks most of our everyday routine interactions with the physical and social world, trust comes into play when routines are disturbed and that pre-reflexive 'taking for granted' is no longer warranted.

And this is precisely why the transition to data-intensive science raises the issue of trust. This transition disrupts routine interactions between scientists and their devices (whether apparatuses or the journals in which they used to publish), between scientists and scientific disciplines, between contributors and users of knowledge, between scientists and citizens, between doctors and patients. As science and technology are collective, cumulative, and material endeavors, people cannot avoid depending on other (past) people and on (past) things. And with this dependence comes vulnerability. This is not a new insight. As historians and social scientists have already noted, science is a social enterprise. A scientist cannot control and check all her colleagues, nor the inner workings of black-boxed scientific devices, and is thus unavoidably vulnerable in the sense that if those colleagues or those devices mess-up, she herself will mess-up too. That is why Steven Shapin says "Knowledge is a collective good. In securing our knowledge we rely upon others, and we cannot dispense with that reliance. That means that the relations in which we have and hold our knowledge have a moral character, and the word I use to indicate that moral relation is trust'. (Shapin 1994, xxv–xxvi).

However, in Luhmann's terms, we should talk about confidence rather than 'trust'—as Shapin does here. In times of what Thomas Kuhn (1962) calls, 'normal', paradigmatic science, awareness of this unavoidable dependence and vulnerability remains in the background. Normal science is typically governed by confidence, and only intermittently by trust. This is due to the fact that science, as a social organization, has developed a host of mechanisms to secure the reliability of knowledge claims. These 'reliability mechanisms'

include research protocols, the checking of instruments, procedures for repeating experiments, regular software updates, instilling a specific ethos in science students (e.g. Merton's universalism, communalism & disinterestedness), institutional guarantees like peer reviewing (cf. Merton's organised skepticism) (Merton 1942), reputation management, and informal networks where trust is grounded in reciprocity and the shared experience of successful collaborations in the past. Thanks to such mechanisms, scientists can have sufficient confidence in their physical and social environment to partake in the collective endeavour of science.

But this changes when a new scientific paradigm shakes up and disrupts existing routines. At such times the established reliability mechanisms, which allowed for confidence rather than trust, no longer function smoothly in the background, but are tested, questioned, rejected, and revised. Rather than supporting the stakeholders of biomedical science, these stakeholders are required to support old and new reliability mechanisms. In this stage of turmoil, the room for confidence decreases, the need for trust increases. This is why the question of digital knowledge and knowledge-making comes up as a question of trust in current healthcare practices.

The five papers that make up this special section explore the trust issues raised by the transition within biomedical science to data-intensive science. They were all written by philosophers and biomedical scientists connected to the Norwegian University of Science and Technology in Trondheim (NTNU). The knowledge repositories that are studied are in most cases (co)built/managed by researchers at the Norwegian University of Science and Technology, and partners.

Rune Nydal, Gaymon Bennett, Martin Kuiper, and Astrid Lægreid (2020) open the special section by pleading for an explicit acknowledgment of the epochal changes effected by the transition to data-intensive science. After extensively mapping the ways in which established ways of doing science are uprooted, they show that many players in the field—rather than acknowledging that this transition entails a profound destabilisation of scientific routines and identities—act as if all is still 'business as usual'. In other words, they cling to an increasingly hollow shell of 'confidence' that frustrates attempts in the community to frankly and constructively discuss the new epistemic risks that characterize the new way of doing science. Because such a discussion is hardly conducted, a healthy situation of reason-based trust is not allowed to grow.

In her contribution, Ane Gabrielsen (2020) argues that data-intensive science may have led to the undermining of traditional mechanisms to enhance mutual trust among scientists, but it has also produced a new type of reliability mechanism. Rather than grounding epistemic trust in individual scientists obeying shared norms, a new discourse of 'open science' now presents openness and transparency as

safeguards for trust in scientific outcomes. However, she shows that this discourse effectively hides the role of an important epistemic community, the data professionals known as biocurators. Biocurators contribute crucially to the trustworthiness of knowledge repositories by safeguarding the quality of the data stored in the repositories. Rather than relying solely on openness and transparency, writes Gabrielsen, data-intensive science needs to acknowledge the role of biocurators in assuring the trustworthiness of data.

In his article ‘Fair trade in building digital knowledge repositories’, Giovanni De Grandis (this issue) explores what would constitute a fair exchange for scientists putting their work and time into biological data repositories. He argues that digital tools are creating new work for these scientists and contributing to ‘science’ or ‘knowledge’ are too abstract goals for overworked and stressed scientists. Rather, for the trade to be fair between knowledge repositories and their contributors, scientists need to be able to trust that their extra work will be rewarding/ed: a concept that De Grandis captures through identifying a new type of trust he calls ‘expediential trust’. Expediential trust is a trust that one’s projects and interests will be expedited, promoted and benefited by engaging in a given project—something that according to De Grandis is contested and contestable once it comes to contributing to digital knowledge repositories.

Lars Ursin, Borgunn Ytterhus, Erik Christensen and John-Arne Skolbekken (2020, this issue) explore another group of stakeholders in biomedical research: participants in biobank research. The article focuses particularly on people who choose to withdraw from this research (‘withdrawers’) comparing them with individuals who chose to continue with the study (‘remainers’). By looking back on interview data collected 16 years ago, just around the birth of the current paradigms of personalised and genomic medicine, Ursin and colleagues examine what conditioned the trust of participants in a large Norwegian, public health study, which was starting to collect the genetic material of existing participants. Ursin and colleagues point out that the worries of participants are still remarkably topical now, and certain topics, such as the risks associated with finding novel uses of participants’ biological materials, worried both remainers and withdrawers. Perhaps one conclusion here is that once it comes to the ethical and social risks associated with biomedical research, the root of these remains shared, as the outcome might be losing one’s autonomy or integrity -contributing to research whose outcomes or uses one might not endorse. What can turn a lack of trust into confidence in biobank research is then a practical but also a philosophical or even psychological question.

Finally, Bjørn Myskja and Kristin Steinsbekk (2020, this issue) apply a Kantian perspective to the trust-relation between doctor and patient. They argue that the transition towards data-intensive healthcare changes that relationship,

as the promise of personalized medicine fits well with the prevalent ideal of patient autonomy. However, personalized medicine will not automatically empower the patient. A major problem is that patients have often great difficulty to make sense of all the new information and lack the necessary expertise to interpret the new information. According to Kant there is both a duty to be trustworthy and a conditional duty to trust others. However, in the new situation it is not so clear who should be trusted. What is therefore needed, according to Myskja and Steinsbekk, are institutional controls that ensure that information is accessible for patients, and medical expertise to support the patients in making sense of that information. Only when those conditions are met, can it be said that active and reflexive trust by the patient is a moral duty.

Together the articles of this special section give an idea of how medical science is entangled with our healthcare system at large, and how the change towards data-intensive science destabilizes established routines inside science, and between science and healthcare practice. They describe how relations between the stakeholders in modern healthcare are changing and how, uncritical, ‘confidence’ is often no longer an option. Instead, ethical work has to be done to recreate relations of trust. The articles describe some of the challenges that have to be met to ensure that under the new and changed circumstances, these stakeholders can—again—trust one another sufficiently to work constructively at realizing the shared goals of our healthcare.

References

- Aronova, E., K.S. Baker, and N. Oreskes. 2010. Big science and big data in biology: from the international geophysical year through the international biological program to the long term ecological research (LTER) Network, 1957—Present. *Historical Studies in the Natural Sciences* 40 (2): 183–224.
- Bok, S. 1978. *Lying: Moral choice in private and public life*. New York: Pantheon Books.
- Baier, A. 1986. Trust and antitrust. *Ethics* 96 (2): 231–260.
- Callebaut, W. 2012. Scientific perspectivism: A philosopher of science’s response to the challenge of big data biology. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 43 (1): 69–80.
- De Grandis, G., and V. Halgunset. 2016. Conceptual and terminological confusion around personalised medicine: a coping strategy. *BMC Medical Ethics* 17 (1): 43.
- De Grandis, G. 2020. Fair trade in building digital knowledge repositories: the knowledge economy as if researchers mattered. *Medicine, Health Care and Philosophy*, <https://doi.org/10.1007/s11019-020-09966-z>.
- EfstathiouNydal, S.R., A. Laegreid, and M. Kuiper. 2019. Scientific knowledge in the age of computation Explicated, computable and manageable? *Theoria Revista de Teoría, Historia y Fundamentos de la Ciencia* 34 (2): 213–236.

- Gabrielsen, A.M. 2020. Openness and trust in data-intensive science: the case of biocuration. *Medicine, Health Care and Philosophy* 23 (3): 497–504.
- Hardwig, J. 1991. The role of trust in knowledge. *The Journal of Philosophy* 88 (12): 693–708.
- Hedgecoe, A. 2004. *The politics of personalised medicine: Pharmacogenetics in the clinic*. Cambridge (UK): Cambridge University Press.
- Hoeyer, K. 2019. Data as promise: Reconfiguring Danish public health through personalized medicine. *Social studies of science* 49 (4): 531–555.
- Kuhn, T.S. 1962. *The structure of scientific revolutions*. Chicago: University of Chicago Press.
- Leonelli, S. 2016. *Data-centric biology: A philosophical study*. Chicago: University of Chicago Press.
- Luhmann, N. (2000) Familiarity, Confidence, Trust: Problems and Alternatives. In *Trust: Making and Breaking Cooperative Relations*, ed. Gambetta, Diego. Electronic edition, Department of Sociology, University of Oxford, chapter 6, pp. 94–107.
- Merton, R.K. 1942. *The Normative Structure of Science [1942] reprinted in Merton 1973*. The sociology of science, Chicago: University of Chicago Press.
- Myskja, B & Steinsbekk, K. (2020) Personalized medicine, digital technology and trust – a Kantian account. *Medicine, Health Care and Philosophy*, this issue.
- Nydal, R., G. Bennett, M. Kuiper, and A. Lægreid. 2020. Silencing trust: confidence and familiarity in re-engineering knowledge infrastructures. *Medicine, Health Care, and Philosophy* 23 (3): 471–484.
- Prainsack, B. 2017. *Personalized medicine: empowered patients in the 21st century?*. New York: NYU Press.
- Shapin, S. 1994. *A social history of truth: Civility and science in seventeenth-century England*. Chicago: University of Chicago Press.
- Ursin, L., Ytterhus, B., Christensen, E., & Skolbekken, J. A. (2020) If you give them your little finger, they'll tear off your entire arm: losing trust in biobank research. *Medicine, Health Care and Philosophy*, this issue.
- Woese, C.R. 2004. A new biology for a new century. *Microbiology and Molecular Biology Reviews* 68 (2): 173–186.
- Wilkinson, M. D., M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos and P. E. Bourne (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data* 3.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.