



OPEN ACCESS

EDITED BY

Yiannis Kyratsis,
Erasmus University Rotterdam, Netherlands

REVIEWED BY

Daniel Simeon - Dubach,
Independent Researcher, Walchwil,
Switzerland
Daniela Capello,
University of Eastern Piedmont, Italy
Io Cheong,
Shanghai Jiao Tong University, China

*CORRESPONDENCE

Michaela Th. Mayrhofer
✉ michaela-th.mayrhofer@i-med.ac.at

RECEIVED 11 May 2025

ACCEPTED 01 August 2025

PUBLISHED 03 September 2025

CITATION

Mayrhofer MT (2025) How the world of
biobanking is changing with artificial
intelligence.
Front. Digit. Health 7:1626833.
doi: 10.3389/fdgth.2025.1626833

COPYRIGHT

© 2025 Mayrhofer. This is an open-access
article distributed under the terms of the
Creative Commons Attribution License (CC
BY). The use, distribution or reproduction in
other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

How the world of biobanking is changing with artificial intelligence

Michaela Th. Mayrhofer^{1,2*}

¹Institute of Human Genetics, Medical University of Innsbruck, Innsbruck, Austria, ²Papillon Pathways e.U., Landskron, Austria

Artificial Intelligence is increasingly shaping the practice of biobanking by influencing how biobanks evolve and operate, especially when it concerns their relationship to data. By assessing four key parameters—size, site, speed, and access—this paper analyzes the impact of AI technologies on biobanks, presenting them as dynamic boundary objects that produce biovalue by transforming biological material and data into intangible assets of the data-driven bioeconomy. Historically rooted at the intersection of health research and healthcare, biobanking is continually reshaped by emerging technologies, policies, and societal expectations. While biobanks were originally defined as collections of samples and associated data, they have recently evolved into complex infrastructures for both data and samples.

KEYWORDS

biobank, artificial intelligence, datafication, health economy, size, site, access, speed

1 Introduction

Biobanking is here. Artificial Intelligence (AI)-driven process automation, data analytics, robotics, the internet, and other rapidly emerging technological advances are driving the revolution of biobanks, biorepositories, and biospecimen science. With the evolution of biobanking from a simple collection of frozen specimens to the virtual biobanks and bioscience seen today, the rise of biobanks brings each nation and its healthcare and economic systems a transformative potential. (1).

The cited passage originates from the exposé titled “Biobanking is Changing the World”, which was published in *Forbes Magazine* about half a decade ago. While many facets of biobanking have been studied extensively throughout the years by various academic disciplines (and continue to be examined), *Forbes Magazine* looked at biobanks with an entrepreneurial lens. Generally, *Forbes Magazine* features trade news and financial information aimed at a target readership in business and technology. As regard to the above excerpt, it is therefore worth noticing that much emphasis is placed on biobanks’ “transformative potential” that is—when unlocked by AI—driving both societal progress and economic prosperity (e.g., “human health” or “nation’s economy”), presenting this as a novelty to the business world.

Rather than being something ‘new’ in and of themselves, biobanks can look back as organized sites that support and satisfy medical curiosity since the 16th century (2). Since a couple of years, biobanks are broadly defined as collections of biological samples and associated data (3), typically located at the intersectoral space between healthcare and health research. Since the late 1990s, the practice of biobanking has gained new meaning—especially for the life sciences—due to standardization,

professionalization and the assemblage of critical mass in both material resources and expertise (4). Consider, moreover, the ISO 20387:2018 definition of biobank as a “legal entity or part of a legal entity that performs biobanking” and biobanking as the “process of acquisition [...] and storing, together with some or all of the activities related to collection, preparation, preservation, testing, analyzing and distributing defined biological material as well as related information and data.”¹ Subsequently, “biobanker” became a distinguished career path that is accredited through post graduate training courses in overall management, data quality, or regulatory and quality aspects.² Today, biobanks can exist either as stand-alone entities³ or integral parts of a clinical infrastructure, especially within university hospitals.⁴ They may be part of transnational networks or research infrastructures (5) and are often considered valuable for private-public partnerships (6), translational research (7) or identified as symbolic locations of national identity (8) or as sites that generate and express bioeconomy (9), and support the interpretation that samples are not isolated objects but also function as data in themselves due to their embedded informational content that is derived through (biomedical) analysis (10).

The *Forbes Magazine* article is now approximately five years old and seems to have aged well. It is thus a good time to revisit its key arguments, especially in relation to the effects that are brought to light through the progression and implementation of AI technologies. As a starting point, it can be noted that countries continue to set up or maintain national biobanks [e.g. (11)]. Second, new technologies such as AI not only shape innovation, but a data-driven world has led to a global race of nations for AI dominance (12). Third, several countries strive to strike for a balance between consumer/citizen/patient rights on the one hand, while enabling open access to data infrastructures and combine data sets to stimulate entrepreneurial innovation on the other hand [e.g. (13)].

The three observations listed above are not comprehensive. Rather, they allow for a suitable argumentative opening to investigate how specific factors have shaped the intersectoral space between healthcare and health research. AI is one of many technologies that affect and shape the practice of biobanking. AI-based algorithms, when applied to biobank data, for instance, can accurately categorize phenotypes by enabling metabolite mapping and demonstrate future clinical applications (14). Alternatively, efficiency is increased by speeding up the analyzation and labeling of images in shorter timeframes by training AI-tools on imaging data from biobanks (15). However,

transformations like these, are not solely driven by technological advancements. They emerge in a regulatory environment and ethical framework that guide their implementation. As such, AI technologies situated in the practice of biobanking serve as a crucial lens through which we can explore the evolving landscape of health data governance, data privacy concerns, and the shifting dynamics between technological innovation and public policy. Thus, following, this brief introduction, the next section of this article discusses the ever contingent and changing practices of biobanks by examining the parameters of size, site, access and speed in relation to AI's potentiality for biobanking.

2 Artificial intelligence: impacting size, site, access and speed in the data-driven health economy

While the collection and categorization of biological materials has developed into an organized practice since the 18th century, computerized databases have supplemented the practice of sample processing and have—since the 1980s—been steadily integrated into the laboratory and scientific work of the biological and biomedical sciences (16, 17). This progression prompted Timothy Lenoir (18) to note that databases challenge laboratories as primary sites of knowledge production, and with it anticipated what Anne Beaulieu called the “informational turn” (19), which describes the turn to data as the now dominant source for scientific knowledge production. Subsequently, data was depicted as a critical resource, exemplarily portraying big data as the “oil of the information economy” (20) or describing national population registers as “goldmines” (21).

Data intense practices share—as noted by many—that they have an unquenchable thirst for ever more data (20, 22). At the same time, especially the health sector is driven by the conviction that datafication will lead to open innovation and precision medicine, including economic growth (23). Snell and others have formulated this development as the turn towards a “regime of data-driven health economy” (24). This regime does not only run on the insatiable thirst for big data, but also on the promise of infinitely commercially exploitable possibilities. Perhaps most crucially, the authors identified the following paradox: namely, that today's system of a data-driven health economy is constructed on the data collection mechanisms of the welfare state, which builds on both the principle of solidarity and a functioning social contract. However, when data extraction is redirected toward private profit and contributes to the erosion of public (health care) systems, the legitimacy of such data practices is called into question. As they argue, “[t]his contradicts the justification of the welfare state data gathering, since a promise of profit itself is not necessarily enough to justify the uses of citizens' personal data as a resource for economic activity and wealth outside data's original, primary context.” (ibid).

Health data today is gathered for many purposes, among them personalized medicine, scientific discovery, disease prevention, lifestyle or self-management of care. The health data industry is a growing sector, and global companies such as Google or Amazon

¹The ISO standard is currently under revision and will be replaced by ISO/CD 20387, see <https://www.iso.org/standard/67888.html> (July 07, 2025).

²E.g., Université Côte d'Azur: <https://univ-cotedazur.eu/msc/biobanks-complex-data-management/career-path> (April 20, 2025).

³E.g., UK Biobank: <https://www.ukbiobank.ac.uk> (April 20, 2025).

⁴E.g., Biobank Graz: <https://biobank.medunigraz.at/en/> (April 20, 2025).

are expanding their ventures into health research and health care (25–27). The employment of health data promises early detection, increased health literacy or tailored therapy. On that ground, the effective utilization of data has become a central objective of health and innovation policies worldwide, including reform plans for national healthcare systems. National strategies seek to harness the potential of health data management in order to enhance (cost- and time-) efficiency in patient care, to enable the stratification of personalized medicine, and to foster research and development—while simultaneously negotiating concerns related to efficiency, fairness, (bio)value, and potential exploitation or discrimination, among other issues (28–31). Moreover, many scholars, while arguing for the employment of AI technologies, do so with caution and call for the study of AI and big data associated risks such as data biases or exclusion criteria (15, 32, 33), especially because data is seen as all-powerful but not as innocent—or as Kelly Bronson described it, “immaculately conceived” (34).

AI algorithms—powered by (big) data—are one of many tools for enabling transformative innovations that influence and shape the practice of biobanking. Due to its perceived “potentiality”, AI has thus become one of the most widely discussed and both morally and economically invested technologies in recent years (71).

In the following sub-chapters, we will employ an AI lens to assess how the practice of biobanking is formed by several key parameters. For this paper, these are size, site, access, and speed. These four parameters serve as the foundation for constructing an analytical framework through which to observe the continuous evolution of a biobank, biobank network or even infrastructure by mobilizing the analytical tools of “boundary objects” (35), “biovalue” (36) and “assetization” (37). Assetization describes how assets are created through both tangible and intangible valuation practices that assign value to resources or concepts and transform them into economic assets. Biovalue refers to the value produced by the biotechnological reformulation of living processes into something else, whereas boundary objects can be defined as entities that operate at the intersection of multiple disciplines and facilitate the meaningful translation of practices across them. Employing these three concepts within the parameters of size, site, access, and speed, ultimately permits describing how biobanks operate at the crossroad between health care and health research by transforming data, samples and new technologies such as AI into assets and therewith generating value for the bioeconomy.

2.1 Size

Let us firstly assess the significance of size by highlighting its importance as a statistical requirement. This is particularly evident in epidemiology, where large datasets are a precondition for any meaningful statistical analysis. Equally so, new research findings necessitate to re-assemble even bigger quantities of fit-for purpose or high-quality datasets. Consequently, quality management standards must be defined, implemented and

checked to ensure, for example, reliability in sample and data analysis, compliance with ethical and legal requirements for (re)use, or clarity regarding data provenance (38, 39). If quality standards are not met, the principle of “garbage-in, garbage-out” (40) would reduce the utility of a biobank not only by volume but by lack of scientific value. In the worst case, a biobank would no longer be considered an asset in the bioeconomic sense (37), and thus without any value at all.

In relation to AI, it is suggested that some well-designed algorithms only need a small, but high-quality dataset to be appropriately trained for purpose, unless it concerns deep learning algorithms that require big data:

The size of the dataset required is directly proportional to the type of AI used and its field of application. Even a large dataset may not be useful if it is noisy, incomplete, or biased. A primary issue is the problem of complex, highly specialized, and specific fields focusing on molecular interactions, protein structures, or drug discovery that typically require domain expertise and specialized knowledge. As a result, the problem space is more constrained, and the available data may be more targeted and focused. In such cases, a smaller sample size can still provide meaningful insights and accurate predictions. (14).

Put differently, data and sample quality are integral aspects of defining the size of what constitutes a “critical mass”. This again requires collaboration practices, which rare disease biobanks were the first to understand:

Another challenge is that RD biobanks need to be connected within networks that ensure uniformly high quality levels of both biomaterials and associated clinical data, apply harmonized operational procedures, reduce redundancies, optimize investments, and facilitate exchanges of expertise and competences. (41).

We have thus established that size matters, but to which degree and why is context dependent and not exclusively linked to statistical power. Paul Burton and others argued that “from a strategic perspective, it is still unclear what ‘large enough’ really means. This question has critical implications for governments, funding agencies, bioscientists and the tax-paying public. Difficult strategic decisions with imposing price tags and important opportunity costs must be taken” (42). Consequently, aspects of size are used for reporting success stories and justifying investments in infrastructure (incl. community) building: “The following biobanks are some of the largest in the world [...]”⁵ or “EBW25 [congress] for the biggest biobank networking opportunity”.⁶ In other words, clinical and national

⁵<https://www.biobanking.com/10-largest-biobanks-in-the-world/> (April 29, 2025).

biobanks have detected that beyond the statistical power of scale, there is also the epistemic power of size that can be employed to funders, customers, or biobankers for creating a sense of pride and purpose. This is equally true when describing the biobank as a national asset and evoking national pride:

When biobanks furthermore acquire a size facilitating claims about national representativity, they potentially come to embody the nation in an almost somatic sense. Biobank freezers can be used metaphorically as a proxy for the surrounding society (I remember once a biobank representative told me that the freezers were where they “kept 80,000 people”), and therefore it is no surprise that large-scale biobanking can become arenas for public negotiation of the duties, entitlements, and mutual obligations between state and citizen. (43).

Ultimately, size goes beyond the quantitative, as Klaus Hoeyer eloquently argues. While size may initially appear to be a neutral, objective, and purely quantitative measure, it is far from an apolitical category. Rather, size often carries embedded assumptions and expectations, strategic implications as well as power dynamics. Consequently, the category of size can be leveraged not only to establish statistical relevance or provide evidence-based justification for a scientific argument, but also to promote a particular collection or biobank strategically, thereby promoting and trading the biobank’s assets, justifying sampling and access strategies or legitimizing the allocation of public funds.

2.2 Site

Let us now consider the category of site which can be physical or virtual, centralized or federated, legalized or a social assemblage. Science and technology studies conceptualize assemblages as relational, experimental, and situated in the context of institutionalized multidisciplinary and/or transnational science collaboration. Assemblages emerge, when certain elements—whether they are material and immaterial in their nature—come together and collectively stabilize a social system such as when a biobank or infrastructure is implemented through discourses, practices, technologies or norms in a specific moment (44). In this sense, biobanks can be described as sites where the materiality of both samples and data is transformed into tangible assets with biovalue, which play a pivotal role for the bioeconomy.

Human genomics is undergoing a step change from being a predominantly research-driven activity to one driven through health care as many countries in Europe now have nascent precision medicine programmes. To maximize the value of the genomic data generated, these data will need to be shared between institutions and across countries. In recognition of this challenge, 21 European countries recently signed a

declaration to transnationally share data on at least 1 million human genomes by 2022. (45).

For expanding on this quote further, the “living organism” metaphor proves useful as it helps depicting the temporal dimension of and spatial organization. It indicates that a biobank, network or infrastructure does not miraculously appear, but is “assembled”, constantly evolving. Just as for an organism, the evolution of a biobank corresponds to the interplay of maturity and social dynamics (46). Biobanks or research infrastructures, in and of themselves—regardless of the sector in which they are implemented—are more than technical constructs. Rather, as Melissa Gilbert and others (47) have argued, they are social systems that play a key role in shaping societal transformations. This is particularly evident when it concerns aspects of equity, inclusion and justice, where the design and conventionalization of infrastructures can either reinforce or obliterate existing disparities. Building on the arguments laid out by Susan Leigh Star and Karen Ruhleder (48), these assemblages of technological and social knowledge might assemble and translate as infrastructures that may be tangible/intangible, federated/centralized, product/process or material/immaterial. Moreover, a village is required, so Christine L. Borgman and Paul Groth (49), to overcome the physical and social distances that lie between data creators, data reusers, data curators or funding agencies. Their understanding goes beyond mere situatedness in a particular context. Rather, they argue that the relational aspects affect the social and technical distances through time and temporality.

We can take from this, that it requires a constant dialogue between the stakeholders to find means to overcome the spatial gap, which is complex and requires infrastructure. “Boundary object,” a concept used in science and technology studies to describe entities that lie at the intersection between communities, is helpful here to localize the space biobanks shape (35). At the same time, infrastructures become more and more complex while striving to sort things out through classification techniques (50). In relation to AI, known aspects such as especially trustworthiness (e.g., human decision making) and economic exploitation (e.g., intellectual property rights or data sovereignty), become more prevalent.

2.3 Access

Let us now turn to the category of access, which is a critical one for any biobank or any data infrastructure as they do not have much value if they cannot be accessed. To manage accessibility, access committees typically define, implement, and monitor the conditions of use, which are a key part of any governance framework and strategy. Consider, for instance the European Union’s ambitious European Health Data Space (EHDS) initiative, which aims to grant citizens increased access to and control of their (electronic) health data across the EU, whilst facilitating health data re-use for public and private research and innovation, as well as policymaking. This initiative clearly positions health data as an asset for primary care but also for research and innovation by defining as one of its key goals to generate public value by increasing access to health data by

⁶<https://www.bbmri-eric.eu/events/europe-biobank-week-2025/> (April 29, 2025).

balancing simplified access for Big Tech innovation without risking the solidarity-based health care systems. Put differently, “the aim of stimulating the European economy by granting free access to citizens’ health data can backfire and have detrimental effects on public trust in and support for medical research” (51). If done wrongly, it can even reinforce digital divides and social inequalities (52). This positions the envisioned national health data access bodies (HDABs) as critical gatekeepers—and possibly bottlenecks—in the implementation of the EHDS. To date, however, HDABs lack a clear and practical mandate, especially because the secondary use of health data derived from electronic health care records is not specifically defined in Article 34.⁷ In addition, access conditions to health data is difficult to harmonize across Europe due to a plethora of diagnostic codes or standards that challenge interoperability:

Even if all EU countries should begin using ICD-11, the same diagnostic codes will be used differently and signify different stages of disease in healthcare systems with different remuneration systems, varying access to healthcare, and diverse registration traditions. Arriving at agreements on semantics and data management procedures among many different stakeholders is a monumental challenge. Indeed, implementation of new EHR [electronic health care record] systems is very time-consuming, costly, and largely beyond healthcare professionals’ remit and scope of action. It is also very challenging, even within a limited (disease or national) area, let alone the entire Union. (51).

This challenge for data sharing across countries is well described in scientific literature, among which the FAIR principles, which stand for “findable, accessible, interoperable, and reusable” (53), are the most prominent ones. Intended as guiding data policy instrument for improved data management across all sciences, the FAIR principles are not only widely accepted in the research community but also heavily promoted by policy makers hoping to stimulate a greater (re)use of health data. To implement FAIR for the practice of biobanking, biological material and data are best conceived as a unified resource. It enables the integration of comprehensive provenance information—including data reuse. It does, however, not provide any indication about fair access. This is not to diminish the importance of technical principles such as FAIR, but rather to emphasize that they are not sufficient on their own to deliver on the complex governance and sensitive data management that is needed to preserve public trust. Enabling fairness requires extensions of the very same principles by additionally promoting data quality, incentivizing data sharing and upholding ethical and privacy preserving practices as suggested by FAIR-health (54). Alternatively, FAIR-er (55) promotes the

inclusion of engagement and participation mechanisms in the design of data governance frameworks. While a lot has been achieved, a high level of complexity remains to make access both FAIR and fair, even in the most advanced countries in health data digitalization, let alone across Europe for the realization of the EHDS (68, 69). Consider, in this context, the example of Findata,⁸ that is designed to grant permits for the secondary use of social and health care data while improving data protection for individuals as a one-stop-shop therewith simplifying access conditions (24).

In the broader context of biobanks and data repositories, the use of AI on health data undoubtedly highlights the interrelated aspects of technology, standards and fair regulation even further. It creates tension between the promotion of AI technologies and data privacy concerns, illustrating how data hunger conflicts with the principle of data minimization (56). Through the deployment of AI, the intertwined aspects of ethical compliance and technical standards become more visible and often require ethical trade-offs inherent to data-intensive practices. “Countries must thus decide how to balance the positive goals of secondary-use activities like healthcare AI with mitigating associated privacy risks. These trade-offs raise issues of resource allocation and justice that have so far been largely neglected in policy debates and the scholarly literature” (57). This especially culminates when access conditions are defined and negotiated, and the authors call for a broader ethical debate on funding priorities rather than just holding AI systems accountable. “This of course requires transparent insight into the available budgets and competing needs. All in all, if such reflections lead to a country explicitly deciding to focus on a strict, conditional or liberal approach to data privacy and/or data access, that decision is morally legitimate if it fulfils conditions of procedural fairness, e.g., accountability and transparency” (57).

Consequently, as many scholars have pointed out, the integration of AI technologies in the healthcare and health research sector must strike a careful balance between free-market forces and open access policies that align with fair (small caps) access, especially when supporting a solidarity-based health system rather than undermining it. For addressing this challenge, digital health strategies need to be developed that are both coherent and in line with the fundamental constitutional values such as rule of law or human dignity. Including such values in the overall data governance and access policy frameworks is a key ingredient for nurturing trustworthiness (24, 57–59). Accordingly, any “boundary object”—such as biobanks situated at the intersection of health care and health research, or AI technologies, increasingly embedded across all sectors of contemporary life—must be aligned with both technical and ethical data access frameworks, integrated within robust governance structures, and, perhaps most importantly—coherently incorporated into the respective health(care) system.

⁷[https://www.european-health-data-space.com/European_Health_Data_Space_Article_34_\(Proposal_3.5.2022\).html](https://www.european-health-data-space.com/European_Health_Data_Space_Article_34_(Proposal_3.5.2022).html) (April 4, 2025).

⁸<https://findata.fi/en/> (April 4, 2025).

Failing to do so risks rendering citizens into commodified objects of an asymmetrical bioeconomy rather than empowered subjects.

2.4 Speed

It is uncontested that innovation in AI is emerging at an exuberant speed. In the world of business and innovation, speed is a category of its own—especially, when it provides a head start and advantages over competitors. While we cannot do full justice on this aspect, it is critical to point out that one, an AI race for global dominance is in full swing and two, it is widely supported by ambitious national strategies and substantial public and private investments (60, 61). Some are in this race to win, whereas others entered the race not to be left behind. In general, as Holzinger and others remind us, “digital transformation can involve the introduction of new technologies and processes to improve the efficiency, accuracy, and speed of research and development and enable the development of entirely new and disruptive products and services” (62).

When looking at the category speed in relation to the integration of AI tools into health research (that embraces innovation by default) and healthcare (that is cautious to change by default), the category of speed is predominantly linked to narratives of efficiency and promises of enhancement of the productivity of research processes and patient care.

Especially for repetitive administrative processes or medical images, AI-assisted tools are expected to expedite, for example, diagnostic workflows by automating repetitive tasks or medical image analysis through faster pattern recognition (70). For biobanking, AI is attributed potentiality (71). It is argued that AI will transform the practice of biobanking even more to the digital space, especially in cancer research, where large datasets available in biobanks are used by machine learning applications that advance the understanding in cancer biology—while building on the decades-long know-how and efficiency of biobanks in sensitive data management (63). At the same time, there is large agreement that it is critical to preserve or build trust(worthiness) in AI systems by retaining accountability through human-in-the-loop or human-in-command approaches. This requires the translation of high-level recommendations into practical processes that can be adhered to regardless the fast pace of technological development and slow regulation (64–67).

3 Conclusion

By examining the categories of speed, site, size and access, this paper explored how AI technologies shape and transform the practice of biobanking, especially in relation to data. For decades, biobanks have been situated at the intersection between health research and healthcare. They have operated as boundary objects that transform the value of samples and data into assets for the bioeconomy. Put differently, biobanks do not merely store samples and data, they actively participate in the co-production of scientific knowledge or governance structures. They shape and are shaped

by their environment. Whereas the practice of biobanking has become more standardized and institutionalized over time, it always was a practice that needed to remain adaptive to new technologies, regulations, societal priorities or national strategies. Biobanks, of late, have transformed into infrastructures that are experienced in engaging with a multitude of stakeholders from within the clinic, the private sector or patient advocacy. Whereas the so-called “data turn” has been unfolding over several decades, AI technologies have nonetheless accelerated the datafication of science and medicine in the last couple of years. Yet, although AI constitutes a significant sociotechnical shift—exemplified in how it reconfigures the parameters of access, site, size, and, above all, speed—it must be situated within the longer trajectory of biobanking as a dynamic and evolving practice.

Author contributions

MM: Writing – original draft, Writing – review & editing.

Funding

The author declares that financial support was received for the research and/or publication of this article. Open Access Funding provided by the Medical University of Innsbruck.

Acknowledgments

The author would like to acknowledge her colleagues of the BBMRI-ERIC ELSI, QM and IT Teams for fruitful and insightful discussions over the years that have given rise to some of the thoughts included in this work, which were, in parts, presented at the BioMed.AI Summer School that took place from 11 to 13 September 2024 in Lisbon, Portugal. The elaborations on size, access, speed and site are solely the author's.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

References

- Pandya J. Biobanking is changing the world. *Forbes*. (2019). Available online at: <https://www.forbes.com/sites/cognitiveworld/2019/08/12/biobanking-is-changing-the-world/> (Accessed April 20, 2025).
- Strasser BJ. Collecting and experimenting: the moral economies of biological research, 1960s–1980s. *Prepr Max-Planck Inst Hist Sci*. (2006) 310:105–23.
- Cambon-Thomsen A. The social and ethical issues of post-genomic human biobanks. *Nat Rev Genet*. (2004) 5:866–73. doi: 10.1038/nrg1473
- Mayrhofer MT. About the new significance and the contingent meaning of biological material and data in biobanks. *Hist Philos Life Sci*. (2013) 35:449–67. Available online at: <https://www.jstor.org/stable/43862195>
- Mayrhofer MT, Holub P, Wutte A, Litton JE. BBMRI-ERIC: the novel gateway to biobanks. From humans to humans. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz*. (2016) 59:379–84. doi: 10.1007/s00103-015-2301-8
- Hämäläinen I, Törnwall O, Simell B, Zatloukal K, Perola M, Van Ommen GB. Role of academic biobanks in public-private partnerships in the European biobanking and BioMolecular resources research infrastructure community. *Biopreserv Biobank*. (2019) 17:46–51. doi: 10.1089/bio.2018.0024
- Murtagh MJ, Demir I, Harris JR, Burton PR. Realizing the promise of population biobanks: a new model for translation. *Hum Genet*. (2011) 130:333–45. doi: 10.1007/s00439-011-1036-3
- Busby H, Martin P. Biobanks, national identity and imagined communities: the case of UK biobank. *Sci Cult (Lond)*. (2006) 15:237–51. doi: 10.1080/09505430600890693
- Mitchell R, Waldby C. National biobanks: clinical labor, risk production, and the creation of biovalue. *Sci Technol Hum Val*. (2010) 35:330–55. doi: 10.1177/0162243909340267
- OECD. *OECD Best Practice Guidelines for Biological Resource Centres*. Paris: OECD Publishing (2007). doi: 10.1787/9789264128767-en
- Tsai Y-Y, Lee W-J. An imagined future community: Taiwan biobank, Taiwanese genome, and nation-building. *BioSocieties*. (2021) 16:88–115. doi: 10.1057/s41292-019-00179-z
- Haecck P. Trump's \$500B AI plan is 'slap' in the face for Europe. *Politico*. (2025). Available online at: <https://www.politico.eu/article/us-500-ai-europe-donald-trump-global-leadership-eu-social-media-china/> (Accessed April 20, 2025).
- Jeng AC, Sibley IJ, Bale TL. A global perspective on AI innovation and effective use in the research lab. *Neuroscience*. (2024) 560:106–8. doi: 10.1016/j.neuroscience.2024.09.034
- Sherlock L, Martin BR, Behsanger S, Mok KH. Application of novel AI-based algorithms to biobank data: uncovering of new features and linear relationships. *Front Med (Lausanne)*. (2023) 10:1162808. doi: 10.3389/fmed.2023.1162808
- Holub P, Müller H, Bil T, Pireddu L, Plass M, Prasser F, et al. Privacy risks of whole-slide image sharing in digital pathology. *Nat Commun*. (2023) 14:2577. doi: 10.1038/s41467-023-37991-y
- Hilgartner S. Biomolecular databases: new communication regimes for biology? *Sci Commun*. (1995) 17:240–63. doi: 10.1177/1075547095017002009
- Hilgartner S. *Reordering Life: Knowledge and Control in the Genomics Revolution*. Cambridge, Massachusetts: The MIT Press (2017).
- Lenoir T. Science and the academy of the 21st century: does their past have a future in an age of computer-mediated networks? In: Voßkamp W, editor. *Ideale Akademie: Vergangene Zukunft Oder Konkrete Utopie?* 11 ed. Berlin: Akademie Verlag (2002). p. 113–29.
- Beaulieu A. From brainbank to database: the informational turn in the study of the brain. *Stud Hist Philos Biol Biomed Sci*. (2004) 35:367–90. doi: 10.1016/j.shpsc.2004.03.011
- MayeR-Schönberger V, Cukier K. *Big Data: A Revolution That Will Transform How we Live, Work, and Think*. Boston: Mariner Books, Houghton Mifflin Harcourt (2014).
- Kongsholm NCH, Christensen ST, Hermann JR, Larsen LA, Minssen T, Pedersen LB, et al. Challenges for the sustainability of university-run biobanks. *Biopreserv Biobank*. (2018) 16:312–21. doi: 10.1089/bio.2018.0054
- Hoeyer K. *Data Paradoxes: The Politics of Intensified Data Sourcing in Contemporary Healthcare*. Infrastructures Series. Cambridge, MA: The MIT Press (2023).
- Hoeyer K. Data as promise: reconfiguring Danish public health through personalized medicine. *Soc Stud Sci*. (2019) 49:531–55. doi: 10.1177/0306312719858697
- Snell K, Tarkkala H, Tupasela A. A solidarity paradox—welfare state data in global health data economy. *Health (London)*. (2023) 27:664–80. doi: 10.1177/13634593211069320
- Neumann K, Mason SM, Farkas K, Santaularia NJ, Ahern J, Riddell CA. Harnessing google health trends data for epidemiologic research. *Am J Epidemiol*. (2023) 192:430–7. doi: 10.1093/aje/kwac171
- Sharon T. The googlization of health research: from disruptive innovation to disruptive ethics. *Per Med*. (2016) 13:563–74. doi: 10.2217/pme-2016-0057
- Slokenberga S. Direct-to-consumer genetic testing: changes in the EU regulatory landscape. *Eur J Health Law*. (2015) 22:463–80. doi: 10.1163/15718093-12341363
- Tupasela A, Snell K, Tarkkala H. The nordic data imaginary. *Big Data Soc*. (2020) 7:2053951720907107. doi: 10.1177/2053951720907107
- Staunton C, Moodley K. Data mining and biological sample exportation from South Africa: a new wave of bioexploitation under the guise of clinical care? *S Afr Med J*. (2016) 106:136–8. doi: 10.7196/SAMJ.2016.v106i2.10248
- Rujano MA, Boiten J-W, Ohmann C, Canham S, Contrino S, David R, et al. Sharing sensitive data in life sciences: an overview of centralized and federated approaches. *Brief Bioinformatics*. (2024) 25:bbae262. doi: 10.1093/bio/bbae262
- Vayena E. Value from health data: European opportunity to catalyse progress in digital health. *Lancet*. (2021) 397:652–3. doi: 10.1016/S0140-6736(21)00203-8
- Fritzsche M-C, Akyüz K, Cano Abadía M, McLennan S, Martinen P, Mayrhofer MT, et al. Ethical layering in AI-driven polygenic risk scores—new complexities, new challenges. *Front Genet*. (2023) 14:1098439. doi: 10.3389/fgene.2023.1098439
- Feldblyum Le Blevenec MK. Selective deployment of AI in healthcare and the problem of declining human expertise. *Bioethics*. (2025). doi: 10.1111/bioe.13424
- Bronson K. *The Immaculate Conception of Data: Agribusiness, Activists, and Their Shared Politics of the Future*. Montreal; Kingston; London; Chicago: McGill-Queen's University Press (2022).
- Star SL, Griesemer JR. Institutional ecology, 'Translations' and boundary objects: amateurs and professionals in Berkeley's museum of vertebrate zoology, 1907–39. *Soc Stud Sci*. (1989) 19:387–420. doi: 10.1177/030631289019003001
- Waldby C. Stem cells, tissue cultures and the production of biovalue. *Interdiscip J Soc Study Health, Illness Med*. (2002) 6:305–23. doi: 10.1177/136345930200600304
- Birch K. Rethinking value in the bio-economy: finance, assetization, and the management of value. *Sci Technol Hum Val*. (2017) 42:460–90. doi: 10.1177/0162243916661633
- Rubeis G. *Ethics of Medical AI*. The International Library of Ethics, Law and Technology, No. 24. Cham: Springer (2024). doi: 10.1007/978-3-031-55744-6
- Zatloukal K, Kungl P. 15 - Biobanks providing a trusted research environment for health data to advance collaborative research and digital transformation of health systems. In: Kumar D, Chadwick R, editors. *Genomics, Populations, and Society*. London: Academic Press (2025). p. 287–94.
- Compton C. Getting to personalized cancer medicine: taking out the garbage. *Cancer*. (2007) 110:1641–3. doi: 10.1002/cncr.22966
- Vaught J, Lockhart NC. The evolution of biobanking best practices. *Clin Chim Acta*. (2012) 413:1569–75. doi: 10.1016/j.cca.2012.04.030
- Burton PR, Hansell AL, Fortier I, Manolio TA, Khoury MJ, Little J, et al. Size matters: just how big is BIG?: quantifying realistic sample size requirements for human genome epidemiology. *Int J Epidemiol*. (2009) 38:263–73. doi: 10.1093/ije/dyn147
- Hoeyer K. Size matters: the ethical, legal, and social issues surrounding large-scale genetic biobank initiatives. *Norsk Epidemiologi*. (2012) 21:211–20. doi: 10.5324/nje.v21i2.1496
- Ong A, Collier SJ. *Global Assemblages: Technology, Politics, and Ethics as Anthropological Problems*. Malden, MA: Blackwell Publishing (2005).
- Saunders G, Baudis M, Becker R, Beltran S, Bérout C, Birney E, et al. Leveraging European infrastructures to access 1 million human genomes by 2022. *Nat Rev Genet*. (2019) 20:693–701. doi: 10.1038/s41576-019-0156-9

46. Akyüz K, Goisauf M, Martin GM, Mayrhofer MT, Antoniou S, Charalambidou G, et al. Risk mapping for better governance in biobanking: the case of biobank.cy. *Front Genet.* (2024) 15:1397156. doi: 10.3389/fgene.2024.1397156
47. Gilbert MR, Eakin H, Mcphearson T. The role of infrastructure in societal transformations. *Curr Opin Environ Sustain.* (2022) 57:101207. doi: 10.1016/j.cosust.2022.101207
48. Star SL, Ruhleder K. Steps toward an ecology of infrastructure: design and access for large information spaces. *Inf Syst Res.* (1996) 7:111–34. doi: 10.1287/isre.7.1.111
49. Borgman CL, Groth P. From data creator to data reuser: distance matters. *Harvard Data Sci Rev.* (2025) 7(2). doi: 10.1162/99608f92.35d32cfc
50. Bowker GC, Star SL. *Sorting Things Out: Classification and Its Consequences.* Cambridge, MA: The MIT Press (2020).
51. Marelli L, Stevens M, Sharon T, Van Hoyweghen I, Boeckhout M, Colussi I, et al. The European health data space: too big to succeed? *Health Policy.* (2023) 135:104861. doi: 10.1016/j.healthpol.2023.104861
52. Van Kessel R, Hrzic R, O'Nuallain E, Weir E, Wong BLH, Anderson M, et al. Digital health paradox: international policy perspectives to address increased health inequalities for people living with disabilities. *J Med Internet Res.* (2022) 24:e33819. doi: 10.2196/33819
53. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR guiding principles for scientific data management and stewardship. *Sci Data.* (2016) 3:160018. doi: 10.1038/sdata.2016.18
54. Holub P, Kohlmayer F, Prasser F, Mayrhofer MT, Schlünder I, Martin GM, et al. Enhancing reuse of data and biological material in medical research: from FAIR to FAIR-health. *Biopreserv Biobank.* (2018) 16:97–105. doi: 10.1089/ib.2017.0110
55. Murtagh M, Machirori M, Gaff C, Blell M, De Vries J, Doerr M, et al. Engaged genomic science produces better and fairer outcomes: an engagement framework for engaging and involving participants, patients and publics in genomics research and healthcare implementation. *Wellcome Open Res.* (2021) 6:311. doi: 10.12688/wellcomeopenres.17233.1
56. Sorell T, Rajpoot N, Verrill C. Ethical issues in computational pathology. *J Med Ethics.* (2022) 48:278–84. doi: 10.1136/medethics-2020-107024
57. Bak M, Madai VI, Fritzsche MC, Mayrhofer MT, McLennan S. You can't have AI both ways: balancing health data privacy and access fairly. *Front Genet.* (2022) 13:929453. doi: 10.3389/fgene.2022.929453
58. De Ruijter A, Hervey T, Prainsack B. Solidarity and trust in European union health governance: three ways forward. *Lancet Reg Health Eur.* (2024) 46:101047. doi: 10.1016/j.lanepe.2024.101047
59. Akyüz K, Cano Abadía M, Goisauf M, Mayrhofer MT. Unlocking the potential of big data and AI in medicine: insights from biobanking. *Front Med (Lausanne).* (2024) 11:1336588. doi: 10.3389/fmed.2024.1336588
60. Teng T. Chinese startup DeepSeek rattles global markets as Nvidia shares plunge. *Euronews.* (2025). Available online at: <https://www.euronews.com/business/2025/01/28/chinese-startup-deepseek-rattles-global-markets-as-nvidia-shares-plunge> (Accessed April 20, 2025).
61. Stargardt G. Details of 110 Billion Euros in investment pledges at France's AI summit. *Reuters.* (2025). Available online at: [https://www.reuters.com/technology/artificial-intelligence/details-110-billion-euros-investment-pledges-frances-ai-summit-2025-02-10/#:~:text=Details%20of%20110%20billion%20euros%20in%20investment%20pledges%20at%20France's%20AI%20summit,-By%20Reuters&text=PARIS%2C%20Feb%2010%20\(Reuters\),billion%20the%20coming%20years](https://www.reuters.com/technology/artificial-intelligence/details-110-billion-euros-investment-pledges-frances-ai-summit-2025-02-10/#:~:text=Details%20of%20110%20billion%20euros%20in%20investment%20pledges%20at%20France's%20AI%20summit,-By%20Reuters&text=PARIS%2C%20Feb%2010%20(Reuters),billion%20the%20coming%20years) (Accessed April 20, 2025).
62. Holzinger A, Keiblinger K, Holub P, Zatloukal K, Muller H. AI for life: trends in artificial intelligence for biotechnology. *N Biotechnol.* (2023) 74:16–24. doi: 10.1016/j.nbt.2023.02.001
63. Frascairelli C, Bonizzi G, Musico CR, Mane E, Cassi C, Guerini Rocco E, et al. Revolutionizing cancer research: the impact of artificial intelligence in digital biobanking. *J Pers Med.* (2023) 13:1390. doi: 10.3390/jpm13091390
64. Lekadir K, Frangi AF, Porras AR, Glocker B, Cintas C, Langlotz CP, et al. FUTURE-AI: international consensus guideline for trustworthy and deployable artificial intelligence in healthcare. *Br Med J.* (2025) 388:e081554. doi: 10.1136/bmj-2024-081554
65. European Commission. Excellence and trust in artificial intelligence (n.y.). Available online at: https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/excellence-and-trust-artificial-intelligence_en (Accessed December 20, 2024).
66. Müller H, Mayrhofer M, Veen E-B, Holzinger A. The ten commandments of ethical medical AI. *Computer (Long Beach Calif).* (2021) 54:119–23. doi: 10.1109/MC.2021.3074263
67. Holzinger A, Zatloukal K, Muller H. Is human oversight to AI systems still possible? *N Biotechnol.* (2024) 85:59–62. doi: 10.1016/j.nbt.2024.12.003
68. De Broe S, Marly N. The European health data space: unlocking the power of health data. *Dev Med Child Neurol.* (2025) 67(8):969–70. doi: 10.1111/dmcn.16359
69. Staunton C, Shabani M, Mascalonzi D, Mezinska S, Slokenberga S. Ethical and social reflections on the proposed European health data space. *Eur J Hum Genet.* (2024) 32:498–505. doi: 10.1038/s41431-024-01543-9
70. KPMG Enterprise. Venture pulse: Q1'18 global analysis of venture funding (2018). Available online at: <https://home.kpmg/xx/en/home/insights/2018/04/venture-pulse-q1-18-global-analysis-of-venture-funding.html> (Accessed April 20, 2025).
71. Mayrhofer MT. Powerful AI. Between accountability and potentiality. In: Mayrhofer MT, Slokenberga S, Mezinska S, editors. *Artificial Intelligence in Biobanking. Ethical, Legal and Societal Challenges.* London: Routledge (2025). p. 206–25. doi: 10.4324/9781032634487