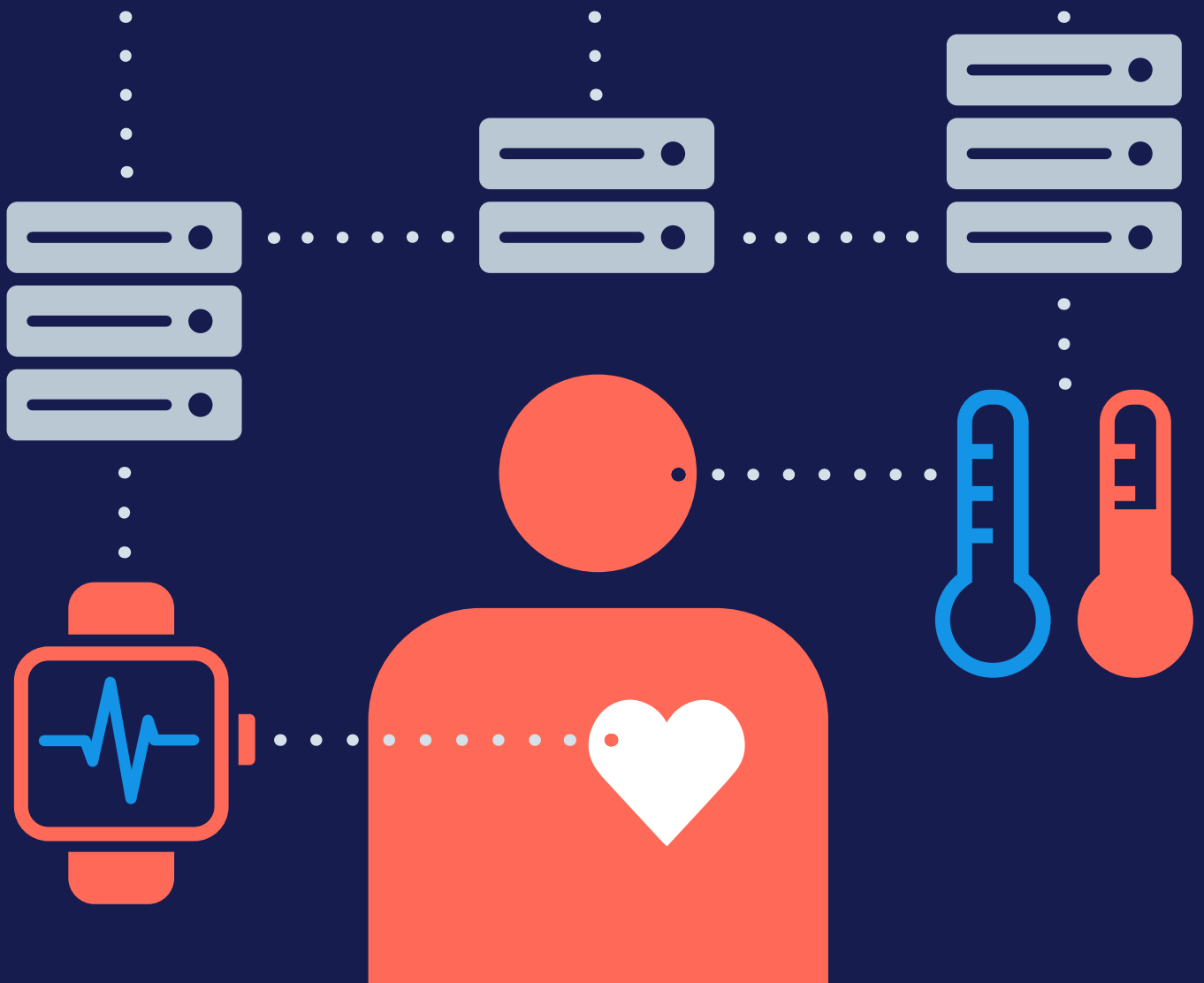




The data will see you now

Datafication and the boundaries of health
October 2020



Contents

- 4** Introduction
- 6** The datafication of health
- 12** Consequences of the datafication of health
- 20** Rethinking health data
- 31** The Internet of Health
- 35** Glossary

Data about people's health is no longer confined to medical records and clinical trials. It is gathered by wearable technologies and smartphone apps; it is inferred from shopping habits and online behaviour; it is used by insurers and advertisers for purposes besides healthcare or medical research. This is the **datafication of health**, and it has significant consequences for people and society.

In this report, the Ada Lovelace Institute explores the datafication of health: what it is, how it occurs, and its impacts on individual and social wellbeing. We draw on examples to synthesise existing research, analyse concepts and surface the societal and ethical challenges arising from the blurring of the boundaries of health data. We chart the mechanisms through which datafication occurs and how people's health becomes '**legible**' to an array of actors, before exploring how datafication challenges existing practical, legal and conceptual definitions of what counts as 'health data'.

We put forward evidence that the datafication of health has created an '**Internet of Health**', in which everyone is made to perform as a patient, subject to continuous monitoring, diagnosis and prediction. We ask what it means to be a patient in this Internet of Health and pose twelve questions that arise from the consequences of datafication.

While we cannot address all the consequences of the datafication of health in full here, we hope that the questions we identify, alongside the concepts explored throughout, act as a foundation for future work by researchers, policymakers, technology developers and health professionals who want to ensure that health data works for people and society.

1. Introduction

The volume of data about health that can be collected from everyday activity has grown exponentially in recent years. Many people have adopted fitness trackers and apps as part of their daily lives, and researchers have linked people's social media and online behaviour data with mental health status. Advertising and retail companies have gathered data about what their customers eat, what fitness services they use and what health products they buy, to make predictions about their health and create targeted marketing.

It may once have been clear that a test performed by a hospital could be used to draw deep insights into someone's health, but that their buying habits could not. Now, the ability to collect datasets, link them and make inferences means that the items someone buys can reveal information about their physical and mental wellbeing. Similarly, where health may once have been discussed only between a patient and a doctor, large technology and data analytics companies now have access to insights about people's heart rate, diet, mental wellbeing and more.

The boundaries of what is or is not data about health have been blurred. What now can count as 'data about health' has expanded to the point where the maxim 'all data is health data' has widespread resonance.¹

As well as the expansion of where and how data about health can be gathered, health is increasingly quantified: information about health is precisely measured, recorded and digitised. Smartwatches don't just monitor the amount of time spent walking, but the exact number of steps and the wearer's heart rate, plotted against time and often linked with geospatial data.

This ever-increasing collection and analysis of quantified data about health has been termed '**datafication**', and it has profound consequences.²

There are considerable benefits to the datafication of health. Not only can individuals be better informed about their health, advances in data analytics lay the foundations for artificial intelligence solutions in diagnosis or drug discovery, as well as for better precision, personalisation and prevention in medicine.

However, datafication raises significant concerns too. It makes individuals' health legible to a broad array of actors outside recognised medical and clinical settings, giving those with the appropriate digital tools an increased ability to know about, and engage with, people's health through their data. Datafication also creates increasingly comprehensive and quantified renderings of health, creating the conditions for disempowerment and providing unprecedented opportunities to monitor and influence people.

1 Warzel, C. (2019) 'All Your Data Is Health Data', *The New York Times*, 13 August. Available at: <https://www.nytimes.com/2019/08/13/opinion/health-data.html> (Accessed: 1 October 2020).

2 Mayer-Schönberger, V, Cukier, K. (2013) *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt.

The datafication of health raises opportunities and challenges that deserve scrutiny. Current legal definitions and practical frameworks for governing health data may no longer be fit for purpose, and the conceptual bounds of 'health data' may need to be rethought.

This report explores datafication and its consequences. It provides the foundations for researchers, policymakers, technology developers and health professionals working with health data to consider the consequences of the datafication of health and ensure that it supports, rather than undermines, individual and social wellbeing.

2. The datafication of health

2.1 What is the datafication of health?

Datafication is the recording of people's everyday life in quantified data.³ In the context of health, datafication is the process where individuals' activity, behaviour and experiences are recorded in quantified data and made analysable by an array of actors in and beyond clinical settings, as reference points for health.

Datafication captures aspects of human activity that can be measured and digitised. An example of this is the step-counting function on the iPhone – part of Apple's Health app.⁴ The iPhone records a person's movements throughout the day, using data from its GPS chip, accelerometer and gyroscope sensors. Through the app, a user can see their activity, visualise their step count and know when they have completed 10,000 steps in a day.⁵

When step-counting, the health benefits are defined not by the quality of the activity itself – walking can be a gentle uphill stroll that stretches the legs, or a strenuous power-walk that works the lungs – but through the number that defines a 'healthy' amount of activity. This can influence how people exercise: those tracking their steps may lengthen their walking times to reach 10,000 steps or shorten their walks once they have completed the required number.

Beyond counting steps, smartphones and wearables can be used to infer far more intimate insights about individuals' health. Modern devices continuously collect a wealth of data that can serve as proxies for individuals' social and physical behaviours.

3 Mayer-Schönberger, V, Cukier, K. (2013).

4 Evans, J. (2018) 'How to turn your iPhone into a step counter', *Computerworld*. Available at: <https://www.computerworld.com/article/3288214/how-to-turn-your-iphone-into-a-step-counter.html> (Accessed: 1 October 2020).

5 Hammond, C. (2019) 'Do we need to walk 10,000 steps a day?', *BBC*. Available at: <https://www.bbc.com/future/article/20190723-10000-steps-a-day-the-right-amount> (Accessed: 1 October 2020).

Researchers in several studies have inferred depression symptoms from passive smartphone data; GPS data can act as a proxy for social withdrawal by measuring how far and for how long a person travels; accelerometer data can be used to analyse physical activity; WiFi-based location data can be used to measure time spent at home or in social environments; data from calendars can be analysed as proxies for stress; and call logs and patterns can be used to analyse social interactions and numbers of contacts.⁶ Using these data points, a Swedish study predicted depression with approximately 60% accuracy and delivered interventions to people displaying depressive symptoms.⁷

By converting aspects of people's lives into quantifiable data, technologies shape the conscious and unconscious choices people make about their health. This creates a 'symbiotic agency' where technologies mediate human experience and affect human agency, while human agency simultaneously shapes the uses of these technologies.⁸ The capacity of data-driven technologies to shape individual choices creates an ongoing need to consider what a 'healthy' relationship between people and their data looks like.

Recording quantified data about health, shaping individuals' experiences of health and giving researchers access to health data in new ways are only a handful of the ways that datafication redefines health and the manifestation of illness in terms of data about people.⁹ Through all these methods, datafication provides new opportunities for the classification, diagnosis and treatment of disease, ill health and wellbeing.

2.2 Inferring and repurposing data about health

The datafication of health relies on the fact that data from devices and online activity can be recomposed and linked with different datasets to draw inferences about people's health. Though most of this data might not be collected for a health purpose, analysing it in a particular way or combining it with another dataset can yield insights about a person's health. Data collected about health can also be adapted for non-health related purposes.

-
- 6 Wahle, F. et al. (2016) 'Mobile Sensing and Support for People With Depression: A Pilot Trial in the Wild', *JMIR mHealth and uHealth*, 4(3), p. e111. doi: 10.2196/mhealth.5960; Onnela, J.-P. and Rauch, S. L. (2016) 'Harnessing Smartphone-Based Digital Phenotyping to Enhance Behavioral and Mental Health', *Neuropsychopharmacology*, 41(7), pp. 1691–1696. doi: 10.1038/npp.2016.7; Aung, M. H., Matthews, M. and Choudhury, T. (2017) 'Sensing behavioral symptoms of mental health and delivering personalized interventions using mobile technologies', *Depression and anxiety*, 34(7), pp. 603–609. doi: 10.1002/da.22646; Renn, B. N. et al. (2018) 'Smartphone-Based Passive Assessment of Mobility in Depression: Challenges and Opportunities', *Mental health and physical activity*, 14, pp. 136–139. doi: 10.1016/j.mhpa.2018.04.003; Ben-Zeev, D. et al. (2015) 'Next-Generation Psychiatric Assessment: Using Smartphone Sensors to Monitor Behavior and Mental Health', *Psychiatric rehabilitation journal*, 38(3), pp. 218–226. doi: 10.1037/prj0000130.
- 7 Wahle, F. et al. (2016) 'Mobile Sensing and Support for People With Depression: A Pilot Trial in the Wild', *JMIR mHealth and uHealth*, 4(3), p. e111. doi: 10.2196/mhealth.5960.
- 8 Neff, G. and Nagy, P. (2018) 'Agency in the Digital Age: Using Symbiotic Agency to Explain Human–Technology Interaction', in Papacharissi, Z. (ed.) *A Networked Self and Human Augmentics, Artificial Intelligence, Sentience*. 1st edn. New York: Routledge, Taylor & Francis Group, 2018. pp. 97–107. doi: 10.4324/9781315202082-8.
- 9 Jain, S. H. et al. (2015a) 'The digital phenotype', *Nature Biotechnology*, 33(5), pp. 462–463. doi: 10.1038/nbt.3223.

For example, sleep quality data from a wearable device might indicate to a person's doctor that their patient has insomnia.¹⁰ But as part of a corporate wellness programme, this same data might be used to track an employee's sleep patterns and to enable an employer to regard the amount and quality of sleep as an employee asset that can be leveraged for improved productivity.¹¹ This data could also potentially be recomposed and used by advertisers who correlate it with data about buying habits to sell the person a targeted sleep programme.¹²

The datafication of health has two defining features, which give those with control over or access to data extensive opportunities to influence people's lives and health:

- **The ability to infer individuals' health status from data**
- **The ability to repurpose data about health for other uses.**

These features in turn shape the datafication of health, giving a multitude of private actors, such as employers, insurers and advertisers an enhanced ability to know about people's health. This creates two areas of particular impact:

- The ability to **infer health status from personal data** has **accelerated the shifting of knowledge about health beyond traditional clinical boundaries** (e.g. from the hospital to the social media analytics company).
- The ability to **repurpose data about health** has meant that **data about health is increasingly used for purposes not directly relating to healthcare.**

10 Persaud, C. (2020) 'Want to track your sleep with your Fitbit? Here's how to do it.' *Android Central*. Available at: <https://www.androidcentral.com/how-track-sleep-fitbit> (Accessed: 28 July 2020).

11 See for example: 'Corporate Sleep Programs'. *Circadian.com*. Available at: <https://www.circadian.com/solutions-services/corporate-sleep-programs.html> (Accessed: 28 July 2020).

12 Sax, M., Helberger, N. and Bol, N. (2018) 'Health as a Means Towards Profitable Ends: mHealth Apps, User Autonomy, and Unfair Commercial Practices', *Journal of Consumer Policy*, 41(2), pp. 103–134. doi: 10.1007/s10603-018-9374-3.

2.3 Datafication through digital phenotyping

One of the key mechanisms that drives the datafication of health is digital phenotyping.

The term '**digital phenotype**' conceptualises the data people produce as an extension of their individual phenotype. A **phenotype** is an individual's biological characteristics, produced by the interaction of their genotype (the collection of genes an individual has) and their environment.¹³ For example, a person's height is a phenotype, and is a result of their genetic inheritance for height expression and the nutrition they receive during childhood.

Biologists have argued that organisms also produce an extended phenotype, which describes the impact they have on their surroundings.¹⁴ Examples include a beaver's dam or a spider's web, which are consequences of the creature's biology and the environment in which they build.

Digital phenotyping was defined in 2015 as 'the moment by moment quantification of the individual level phenotype in situ using data from personal digital devices'.¹⁵ In other words, digital phenotyping is the measurement and analysis of people's interactions with networked digital technologies like smartphones, social media apps, loyalty cards, websites and more.¹⁶

The traces people leave behind when they interact with networked digital technologies can be used to understand their activity, behaviour and other characteristics, yielding insights about their health. Digital phenotyping blurs the perceived boundaries of the body, meaning it is possible to make algorithmically driven inferences about health based on disembodied data about people.

Using data from digital phenotyping to understand health reflects the latest in a line of epistemic shifts in medicine, driven by developments that have transformed and made visible new representations of the body. It is comparable to the X-ray machine giving clinicians 'new sight' in the late 1800s, by revealing a 'hidden existence' that made broken bones legible through visual images.¹⁷

Digital phenotyping similarly reveals a 'hidden existence' that redefines the manifestation of illness through its representation in data and offers a 'new sight'. It provides novel ways to monitor, diagnose and treat disease in real time, beyond traditional clinical boundaries.

13 For more information on phenotypes, see: Scitable. phenotype / phenotypes. Available at: <https://www.nature.com/scitable/definition/phenotype-phenotypes-35/> (Accessed: 1 October 2020). For genotypes, see: 'The Human Genome Project.' Genome.gov. Available at: <https://www.genome.gov/human-genome-project> (Accessed: 28 July 2020).

14 Dawkins, R. (2016) *The Extended Phenotype: The Long Reach of the Gene*. Oxford, New York: Oxford University Press (Oxford Landmark Science).

15 Torous, J. et al. (2016) 'New Tools for New Research in Psychiatry: A Scalable and Customizable Platform to Empower Data Driven Smartphone Research', *JMIR mental health*, 3(2), p. e16. doi: 10.2196/mental.5165.

16 Jain, S. H. et al. (2015a). 'The digital phenotype'

17 See: The Photogram, Volume 5 (1898) *Princeton University*. Available at: https://books.google.co.uk/books/about/The_Photogram.html?id=bFoXAAAAYAAJ&redir_esc=y (Accessed: 6 October 2020).

This means information about health no longer needs to be communicated between a patient and their doctor, or produced through tests in a clinical lab. Digital phenotyping and the datafication of health mean that data about health can be inferred from people's everyday existence in the digital world. Further, because digital phenotyping takes place using everyday technology rather than specialised, calibrated equipment in medical environments, the diagnostic accuracy of digital phenotyping has dependencies on whether collected data is reliable and accurate.

2.4 The increasing datafication of health

Digital phenotyping promises new and more efficient ways to diagnose and prevent disease. Researchers have found that AI has helped diagnose depression three months earlier than health services by analysing Facebook posts, and in 2017 Facebook rolled out an algorithm that detects the likelihood that an individual is at an increased risk of self-harm.^{18,19} Assuming the algorithm was accurate and free from biases, and if this data could be shared securely with researchers and mental health professionals, these algorithmic insights could play a valuable role in early intervention and mitigation of mental illness in the future.

However, there are also significant concerns around the inferring of sensitive personal information by online platforms. When Facebook makes inferences about individuals' health from 'non-medical data sources', it does not classify the information it produces as sensitive health information.

Many people are more likely to feel comfortable giving their doctor intimate information about their health because of the safeguards that prevent this information being transferred beyond clearly defined medical boundaries, like patient-doctor confidentiality and the ethical codes of conduct followed by health professionals. But the datafication of health fundamentally challenges people's privacy expectations, by enabling data about health to flow in new ways beyond previously established conceptual, contextual and regulatory boundaries.

Safeguards protect privacy by maintaining the 'contextual integrity' of information; in medical settings, assurances for patient privacy are established and implicit.²⁰ But the datafication of health means it is possible for multiple actors across different contexts, such as insurance, advertising and employment, to infer and repurpose aspects of an individual's health status without the safeguards present in medical settings. Current data practices mean that these inferences could be made about an individual without their knowledge, let alone their consent.

18 Eichstaedt, J. C. et al. (2018) 'Facebook language predicts depression in medical records', *Proceedings of the National Academy of Sciences of the United States of America*, 115(44), pp. 11203–11208. doi: 10.1073/pnas.1802331115.

19 Goggin, B. (2019) 'Inside Facebook's suicide algorithm: Here's how the company uses artificial intelligence to predict your mental state from your posts', *Business Insider*. Available at: <https://www.businessinsider.com/facebook-is-using-ai-to-try-to-predict-if-youre-suicidal-2018-12> (Accessed: 1 October 2020).

20 Nissenbaum, H. (2009). *Privacy in context: Technology, policy, and the integrity of social life*. Stanford CA: Stanford University Press.

The algorithmic processing of people’s health data promises new opportunities for managing individual health on a population-wide scale. However, it also offers the opportunity for unparalleled exercise of power over the governance of people’s bodies and lives. People’s interactions with online services and networked devices have blurred traditional conceptions of health data, to a point where the datafication of health challenges fundamental ideas about what constitutes health data in practical and legal terms. Potentially, all the data we produce is now data about our health.²¹

Increasing datafication of health

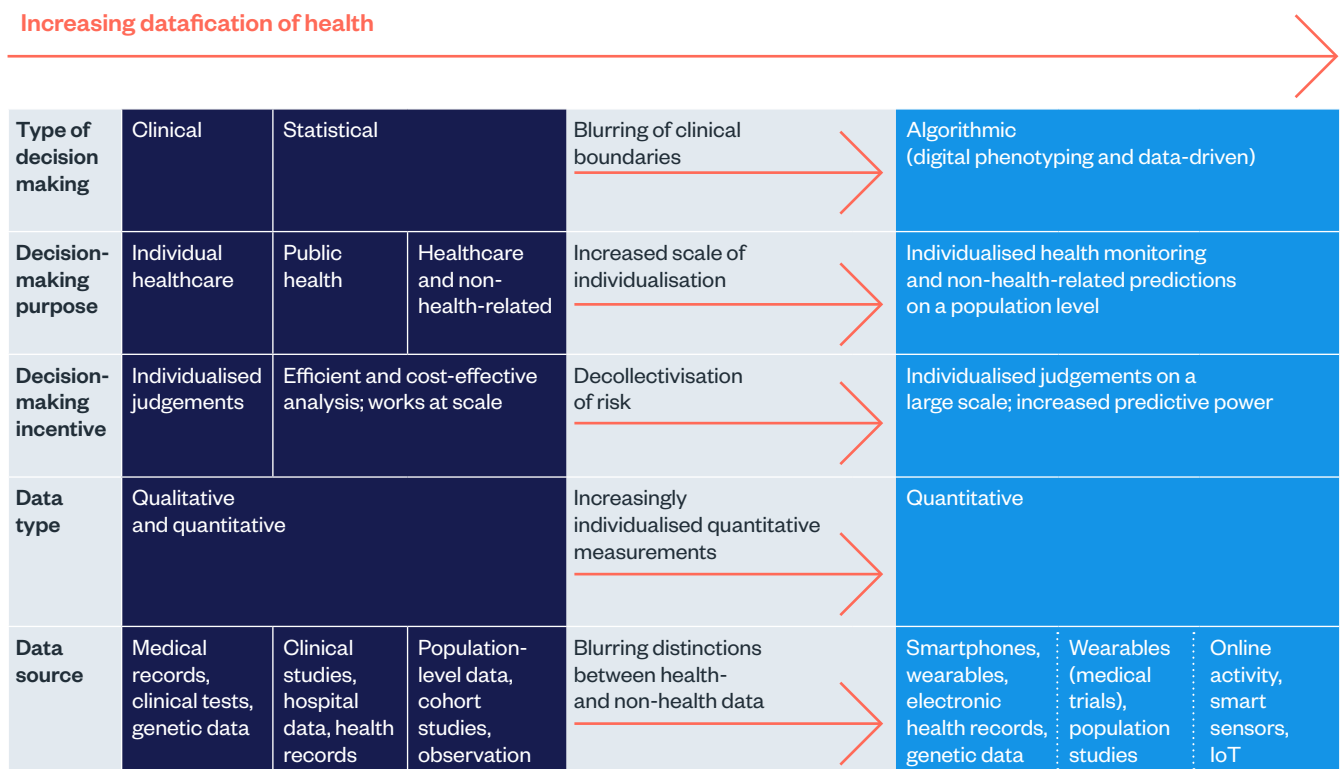


Figure 1
The increasing datafication of health influences healthcare delivery and blurs boundaries

21 Sobhani, M. (2019) *HIPAA Isn’t Enough: All Our Data is Health Data*, USC Center for Body Computing. Available at: <https://www.uscbodycomputing.org/uncensored/hipaaisnotenough> (Accessed: 1 October 2020).

3. Consequences of the datafication of health

3.1 Legibility and the private sector

The increasing datafication of health has not only expanded what counts as data related to health. It has also changed the contexts in which data about health is generated gathered, stored and processed.

What does it mean to be legible?

By making human activity analysable in terms of health, datafication has made an increasing range of human experience legible.

Legibility, as defined by Linnet Taylor, is 'the way in which citizens become visible, or, legible, to authorities [public or private] through data collection and analysis'.

The datafication of health provides actors who have access to the appropriate data and technology with an enhanced ability to know about health – albeit datafied representations of health. As such, datafication makes citizens and their health legible to a broader array of actors than before, and begins to redefine the scope of health within public life.²²

Previously, because of the limited availability of medical technologies and the specialised knowledge required to produce meaningful insights about health, data about health was primarily collected and analysed within strict regulatory boundaries, such as hospital testing labs or approved clinical trials.

Now, data about people's health can exist increasingly within privately owned networks and digital infrastructures, making a person's health status legible to new and different actors.

Within these infrastructures, health data is extracted and used in a way that means private actors can make inferences about people's health and repurpose data about it for their own ends. Where people have become legible to private actors, the datafication of health not only shapes what counts as data about health, but changes how people's health can have value, and who can benefit from that value.

This feature of data about health is a product of the increasing significance of data in a world mediated by online platforms. When early mobile health technologies emerged in the 2000s, 'wireless telemedicine' systems were designed as closed loops that linked a patient to a clinician via dedicated hospital or medical servers.²³

22 See: Taylor, L. (2016) *Data subjects or data citizens? Addressing the global regulatory challenge of big data*, *Information, Freedom and Property*. Routledge. doi: 10.4324/9781315618265-13.

23 Pattichis, C. et al. (2002) 'Wireless Telemedicine Systems: An Overview', *Antennas and Propagation Magazine, IEEE*, pp. 143–153. doi: 10.1109/MAP.2002.1003651.

Now, modern networks that support mobile devices, online platforms and the Internet of Things (IoT)²⁴ mean that the datafication of health no longer occurs solely within closed loops focused on patient experience and delivery of care.

Now, datafication is shaped by incentives to gather personal data of all kinds to make profitable predictions about people. As Greenwood et al. explain, the ‘digital breadcrumbs we leave behind are clues to who we are, what we do, and what we want. This makes personal data – data about individuals – immensely valuable, both for public good and for private companies.’²⁵

The datafication of health is valuable within the modern political economy of data in two key respects: data about health empowers the actors who have jurisdiction over the data with a greater ability to understand, predict and mitigate risk, and it gives these actors an enhanced authority with which to profile and categorise individuals on this basis.

1. Data about health provides granular insights into human activities and experiences, which can be used to strengthen behavioural predictions

The datafication of health is intimately tied up with imperatives in the economy of behavioural prediction and surveillance capitalism, where vast amounts of granular data about people are fed into systems that make predictions, target advertising and drive services.²⁶

In this context, inferring data about health provides a valuable resource of data about people’s online activity (through their digital phenotype), which can be fed through machine learning algorithms to generate valuable predictions – such as what type of adverts might resonate with them, or what their credit score might be.²⁷

This perspective makes clear the strong economic incentives underpinning the development of technologies that facilitate the datafication of health. Advertisers see wearables as a source of granular insights and new behavioural data related to users’ activity, health and emotional state.²⁸

-
- 24 Burgess, M. (2018) ‘What is the Internet of Things? WIRED explains’, *Wired UK*, 16 February. Available at: <https://www.wired.co.uk/article/internet-of-things-what-is-explained-iot> (Accessed: 28 July 2020).
- 25 Greenwood, D, et al. (2014). ‘The New Deal on Data: A Framework for Institutional Controls’, In J. Lane, V. Stodden, S. Bender, & H. Nissenbaum (Eds.), *Privacy, Big Data, and the Public Good: Frameworks for Engagement* (pp. 192-210). Cambridge: Cambridge University Press. doi:10.1017/CBO9781107590205.012.
- 26 Zuboff, S. (2018). *The age of surveillance capitalism: the fight for the future at the new frontier of power*. London, Profile Books.
- 27 Naughton, J. (2019) “‘The goal is to automate us’: welcome to the age of surveillance capitalism”, *The Observer*, 20 January. Available at: <https://www.theguardian.com/technology/2019/jan/20/shoshana-zuboff-age-of-surveillance-capitalism-google-facebook> (Accessed: 1 October 2020).
- 28 Ovum (2016) Digital marketing strategies to get ahead in the high-tech industry. Available at: <https://landing.adobe.com/dam/2016/downloads/341847.en.high-tech-report.pdf> (Accessed: 1 Oct 2020).

Google's acquisition of Fitbit exemplifies this incentive; along with Fitbit's hardware, Google will acquire the activity data of 28 million users.²⁹ Google – or Alphabet, its parent company – has already curated the most extensive population datasets in existence through its search engine user data. The addition of fitness and activity data would potentially provide unprecedented insights into individuals' health.

Even assuming the most positive intentions for those who have jurisdiction over this data, questions are raised by the sheer power and influence this scale and scope of insight will enable Alphabet, and other companies who have access to data about people's health, to have over people's health and lives.

Without adequate legal protection circumscribing the boundaries and uses of this data, people will be disempowered when decisions and inferences are made about their health that they cannot question, challenge, understand or know about. An individual might find they have been denied life insurance on the grounds that they are at high risk of heart disease, but this risk may have been indicated from data from their digital phenotype rather than from sources they know about, such as their medical notes.

2. Knowledge about an individual's health status gives actors' predictions an enhanced authority, bolstering practices of scoring, categorising and profiling

Knowledge about people's health gives actors including companies, governments and organisations biopolitical power – the power to have greater influence over the governance of people's bodies. Those with knowledge about individuals' health can use biopolitical categorisations, such as 'healthy' vs 'ill' or 'normal' vs 'abnormal', to discriminate and differentiate more decisively between people.³⁰ For example, financial technology companies ('fintechs') can make psychological measurements and assessments to reveal traits that can provide a basis to prevent individuals from accessing a loan.³¹

The medicalised nature of this knowledge mobilises the normative authority of medicine. Using this authority, inferred data about people's health can be used to bolster practices of ordering, categorising and profiling individuals, conferring greater credibility on the outcomes of those practices because of their medical basis. Where these authoritative predictions are used in contexts of risk analysis, such as credit scoring, insurance, advertising and corporate wellness, this raises important questions about the future of individual privacy, agency and empowerment.

29 Gartenberg, C. (2019) 'Google buys Fitbit for \$2.1 billion', *The Verge*. Available at: <https://www.theverge.com/2019/11/1/20943318/google-fitbit-acquisition-fitness-tracker-announcement> (Accessed: 1 October 2020); Sawh, M. (2019) '5 reasons why Google just bought Fitbit', *TechRadar*. Available at: <https://www.techradar.com/news/5-likely-reasons-why-google-just-bought-fitbit> (Accessed: 28 July 2020).

30 Ajana, B. (2017) 'Digital health and the biopolitics of the Quantified Self', *Digital Health*, 3, p. 2055207616689509. doi: 10.1177/2055207616689509.

31 Vasagar, J (2016) 'Kreditech: A credit check by social media.' *Financial Times*. Available at: <https://www.ft.com/content/12dc4cda-ae59-11e5-b955-1a1d298b6250> (Accessed: 12 October 2020).

For example, the potential for native advertising in mobile health (m-health) technologies, like wearables and smartphones, provides examples of how technology companies can use data to merge health and commercial content in ways that are hard to detect.^{32,33}

The GPS technology company Garmin sells a range of wearables that are integrated with the 'Garmin Connect' app.³⁴ The app provides tailored digital insights based on a user's recorded statistics. One of these insights is to 'learn good sleep habits', and embedded within the insight is a link promoting a sleep programme that can be purchased for \$40. Another fitness company, Under Armour (UA), has acquired multiple major m-health apps and integrated them into 'the world's first 24/7 connected health and fitness system', linking with UA Healthbox devices.³⁵ UA's Chief Digital Officer has emphasised the company's desire to link the apps more closely, so that after tracking someone's run MyFitnessPal could suggest a place for a runner to purchase a snack.³⁶

These examples highlight the concern that the personalisation of m-health apps increases their persuasive, commercial potential.³⁷ This is amplified by their inferred medical authority, and legitimised through the idea that personalisation equals improved user experience. Karen Yeung describes techniques where big data interacts with personalisation as '**hypernudges**': highly persuasive and personalised influences on people's behaviour that are increasingly hard to detect or counter.³⁸

Many of these devices and applications are also emerging in corporate settings, especially in countries like the USA, where health insurance costs fall on employers, and potential cost savings have similarly incentivised the incorporation of big data practices. These trends enable big tech to reimagine and restructure areas of the established healthcare market. For example, health tech companies such as Fitbit and Jawbone have launched enterprise solutions, selling trackers as well as analytics software to employers, to help them assess and incentivise workplace wellness.³⁹

-
- 32 For a definition of Native Advertising, see: Sharethrough. Native Advertising - The Official Definition. Available at: <https://www.sharethrough.com/nativeadvertising/> (Accessed: 1 October 2020).
- 33 Sax, M., Helberger, N. and Bol, N. (2018). 'Health as a Means Towards Profitable Ends'.
- 34 Garmin Connect™ – Apps on Google Play. Available at: https://play.google.com/store/apps/details?id=com.garmin.android.apps.connectmobile&hl=en_GB (Accessed: 28 July 2020).
- 35 See: 'Under Armour Record' (no date) *Download.com*. Available at: https://download.cnet.com/Under-Armour-Record/3000-2129_4-78279369.html (Accessed: 28 July 2020); 'UA HealthBox: Your Connected Fitness System' (no date) Under Armour. Available at: <https://www.underarmour.co.uk/en-gb/healthbox-your-connected-fitness-system.html> (Accessed: 28 July 2020).
- 36 Kell, J. (2016) 'How Under Armour wants to use data to make you healthier.' *Fortune*. Available at: <https://fortune.com/2016/04/05/under-armour-apps-healthier/> (Accessed: 28 July 2020).
- 37 Kant, T. (2020) *Making it Personal: Algorithmic Personalization, Identity, and Everyday Life*. Oxford, New York: Oxford University Press.
- 38 Yeung, K. (2016) 'Hypernudge': Big Data as a Mode of Regulation by Design'. *SSRN Scholarly Paper* ID 2807574. Rochester, NY: Social Science Research Network. Available at: <https://papers.ssrn.com/abstract=2807574> (Accessed: 1 October 2020).
- 39 Fitbit. One Easy Health Solution for All Your Employees. Available at: <https://healthsolutions.fitbit.com/> (Accessed: 12 October 2020); Allison, C. (2019) Jawbone Health: Everything we know so far about the new premium health platform, *Wearable*. Available at: <https://www.wearable.com/health-and-wellbeing/jawbone-health-hub-medical-health-wearable-2144> (Accessed: 12 October 2020).

A study by the Harvard School of Public Health found that companies save more than \$3 for every \$1 spent on wellness programmes.⁴⁰ But a further study found that those savings are difficult to achieve without being discriminatory toward workers with high health risks, as savings to employers may come from shifting costs to those with higher health burdens. This means the most vulnerable may, in effect, subsidise healthier colleagues.⁴¹

Alongside wearables companies, a host of emerging start-ups, such as Jiff and Wellbe, are building customised, mobile and web experiences that connect company incentive and benefit programmes to 'healthy employee behaviours' using wearables and health apps.⁴² The measurement of employee activity through wearable devices not only demonstrates an increased medicalisation of life beyond clinical boundaries, but also involves the integration of lifestyle and labour, where employees still have obligations beyond the workplace to maintain the 'healthy lifestyle' prescribed by their employer.⁴³

These kinds of data-driven recommendations made by private-sector technology companies are not subject to the same oversight and accountability mechanisms that apply to medical recommendations made in traditional healthcare contexts, such as ethical codes and practices adhered to by doctors, or strict legal duties for health professionals.

Robust accountability empowers patients and maintains standards of care, patient safety and access to legal recourse. If a patient follows a doctors' advice but falls seriously ill, the doctor can be held accountable. But if a user follows the data-driven recommendations and nudges of a dietary and fitness application, then over-exercises, under-eats and falls seriously ill, it is unclear who is liable or what recourse they might have.

The changing contexts within which data about health can be gathered, stored and processed is beginning to allow new actors, beyond the hospital or doctor's office, to influence an individual's experience of health and care. Private actors can now promote their commercial ends using data-driven recommendations about health, blurring the boundaries between health and lifestyle, and patient and consumer.

40 Baicker, K., Cutler, D. and Song, Z. (2010) 'Workplace Wellness Programs Can Generate Savings', *Health Affairs*, 29(2), pp. 304–311. doi: 10.1377/hlthaff.2009.0626.

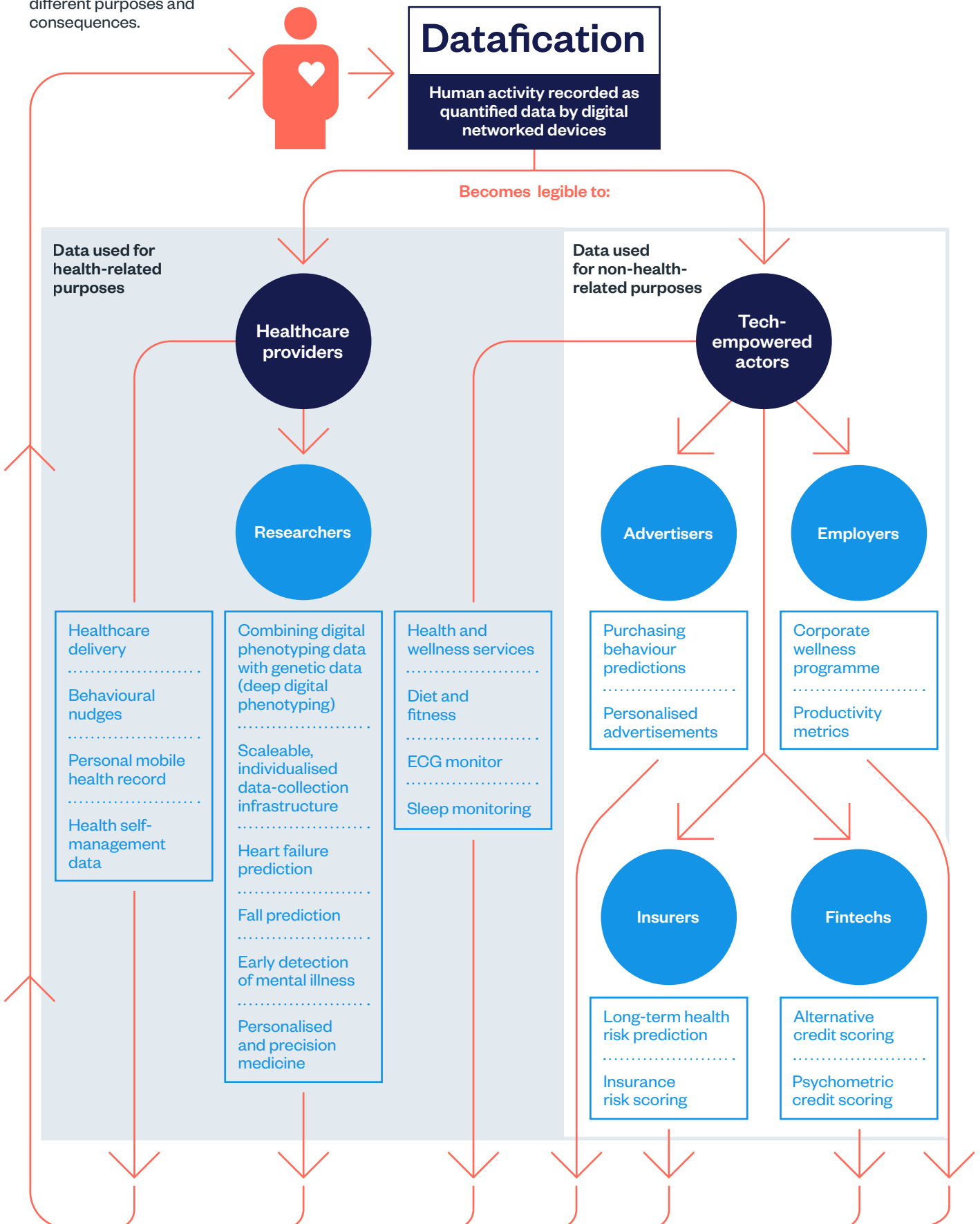
41 Horwitz, J. R., Kelly, B. D. and DiNardo, J. E. (2013) 'Wellness Incentives In The Workplace: Cost Savings Through Cost Shifting To Unhealthy Workers', *Health Affairs*, 32(3), pp. 468–476. doi: 10.1377/hlthaff.2012.0683.

42 See: 'Jiff - Crunchbase Company Profile & Funding.' *Crunchbase*. Available at: <https://www.crunchbase.com/organization/jiff> (Accessed: 6 October 2020); 'Welbe - Crunchbase Company Profile & Funding' *Crunchbase*. Available at: <https://www.crunchbase.com/organization/welbe> (Accessed: 6 October 2020).

43 Ajana, B. (2017) 'Digital health and the biopolitics of the Quantified Self', *Digital Health*, 3, p. 2055207616689509. doi: 10.1177/2055207616689509.

Figure 2

The datafication of health makes individuals legible to different actors in different ways, and with different purposes and consequences.



3.2 Legibility and clinical care

In the provision of healthcare, the twin incentives of ‘better health outcomes’ and ‘lower health costs’ are powerful shaping forces.⁴⁴ The datafication of health can play a key role in realising both incentives by enabling patient monitoring and engagement beyond traditional clinical boundaries.

In creating better health outcomes, the datafication of health is instrumental in the emergent paradigm of ‘**P4 medicine**’, which outlines a vision of healthcare that is predictive, preventative, personalised and participatory. This vision emphasises a shift away from reactive medicine, which is focused on treating established disease, to a more proactive practice, focused on disease prevention, health management and wellness optimisation.⁴⁵

Realising this vision will rely on the datafication of health and practices such as digital phenotyping that consider an individual’s health in relation to their environment, activity and lifestyle.

Visions of datafied healthcare promise to improve health outcomes and reduce costs. In high income countries, the combination of chronic disease and ageing populations incentivises cost-cutting measures and promotes the move towards care systems where the value of care is derived from measuring health outcomes against their cost.⁴⁶ Both the management of chronic disease and the implementation of value-based care benefit from tech-based solutions.

Big data can be used to identify and manage high-cost patients, predict and prevent readmissions, triage patients effectively and organise clinical workflows.⁴⁷ Data-driven mobile devices also play a key role in reducing costs to health services through improved patient engagement and **activation**, increasing the role patients play in their own care and improving their understanding of the care process.⁴⁸ In 2014, the European Commission emphasised the potential of m-health ‘to play a part in the transformation of healthcare and increase its quality and efficiency’.⁴⁹

Across the world, health spending is equivalent to approximately 10% of global GDP.⁵⁰ The healthcare market is a highly lucrative area, and the desire to improve efficiency in care draws convergence between established healthcare providers wanting to modernise patient care and tech companies looking to establish themselves further within the market.

-
- 44 Dentzer, S. (2013) ‘Rx For The “Blockbuster Drug” Of Patient Engagement’, *Health Affairs*, 32(2), pp. 202–202. doi: 10.1377/hlthaff.2013.0037.
- 45 Flores, M. et al. (2013) ‘P4 medicine: how systems medicine will transform the healthcare sector and society’, *Personalized Medicine*, 10(6), pp. 565–576. doi: 10.2217/pme.13.57.
- 46 ‘What Is Value-Based Healthcare?’, *NEJM Catalyst*. (2017) Available at: <https://catalyst.nejm.org/doi/abs/10.1056/CAT.17.0558> (Accessed: 6 October 2020).
- 47 Brigham and Women’s Hospital. (2014) Six cases where big data can reduce healthcare costs. *ScienceDaily*. Available at: <https://www.sciencedaily.com/releases/2014/07/140708165813.htm> (Accessed: 28 July 2020).
- 48 Dentzer, S. (2013)
- 49 European Commission (2014) *Green Paper on mobile health (‘mHealth’), Shaping Europe’s digital future*. Available at: <https://ec.europa.eu/digital-single-market/en/news/green-paper-mobile-health-mhealth> (Accessed: 6 October 2020).
- 50 The World Bank (2020) Current health expenditure (% of GDP). Available at: <https://data.worldbank.org/indicator/SH.XPD.CHEX.GD.ZS> (Accessed: 6 October 2020).

Electronic personal health records are one example where these motivations converge, and they have been described as necessary to improving patient care and for patient empowerment.⁵¹ While many electronic health records systems are developed solely for use in clinical settings, there is significant potential to merge and link data from other settings too.

Tech giant Apple, through the Apple Health Record platform, is connecting patients in a growing number of health institutions with its medical records.⁵² It is also building an ecosystem that connects digital health and wellness apps with third-party researchers and developers in the same service.⁵³ The Apple Watch, a key hardware component of this ecosystem, dominates the global wearable market.⁵⁴ These developments are crucial to Apple's quest to deliver digital services, but further locks users into the Apple ecosystem, and offers Apple ever-more influence over aspects of people's lives.

Datafication and m-health technologies have the potential to be most disruptive in nations that lack physical healthcare infrastructure and have large, underserved rural communities.⁵⁵ In these settings, the proliferation of smartphones accelerates the datafication of health, and could establish datafied health as the norm in healthcare delivery.

Mobile health company Babylon has announced a 10-year partnership with the Government of Rwanda, aiming to make Rwanda 'the most advanced country for digital health'.⁵⁶ But the opportunity to provide quality care in these settings comes with the cost of increasing influence and power over citizens and governments by those who provide and manage the digital and data infrastructure.

The datafication of health is shaped by positive incentives to reduce costs and improve care. But the increasing involvement of private technological infrastructures in data-driven approaches to healthcare means that non-clinical incentives are also appearing.

These developments blur the boundaries where data about people's health is collected, stored and processed, and who has access to it. This poses urgent challenges for the way that health data is practically and legally governed.

-
- 51 Alhomod, A. S. and Alzahrani, S. (2019) 'Patient Empowerment via Mobile Personal Health Records and Mobile Health Applications: A Review of the Current Use', *2019 Portland International Conference on Management of Engineering and Technology (PICMET)*. pp. 1–4. doi: 10.23919/PICMET.2019.8893847.
- 52 Apple (2018) 'Doctors put patients in charge with Apple's Health Records feature.' *Apple Newsroom*. Available at: <https://www.apple.com/newsroom/2018/03/doctors-put-patients-in-charge-with-apples-health-records-feature/> (Accessed: 28 July 2020).
- 53 CB Insight (2019) 'Apple Healthcare: Health Plan & Strategy'. *CB Insights Research*. Available at: <https://www.cbinsights.com/research/apple-healthcare-strategy-apps/> (Accessed: 28 July 2020).
- 54 IDC (2020) *Earwear and Wristbands Drive First Quarter Growth in the Worldwide Wearables Market*. Available at: <https://www.idc.com/getdoc.jsp?containerId=prUS46432620> (Accessed: 6 October 2020).
- 55 Latif, S. et al. (2017) 'Mobile Health in the Developing World: Review of Literature and Lessons From a Case Study', *IEEE Access*, 5, pp. 11540–11556. doi: 10.1109/ACCESS.2017.2710800.
- 56 Babylon Health (2020) 'Rwanda will be the world's most advanced country for digital health', Available at: <https://www.babylonhealth.com/blog/business/rwanda-will-be-the-worlds-most-advanced-country-for-digital-health> (Accessed: 6 October 2020).

4. Rethinking health data

4.1 Defining health data

Effective definitions of health data are crucial to the proper governance, protection and use of health data – both to protect individuals but also to unlock the potential of data to improve health. However, the question ‘What is health data?’ prompts different answers depending on how the scope of the category ‘health data’ is interpreted.

A narrow practical interpretation of health data is data collected within the running of a healthcare system, for example in the delivery of healthcare services or the conducting of medical research. But broader definitions consider any data relating to health, such as any data relating to the physical or mental health of an individual. This could include data about environmental or social factors that affect health.

These definitions of health data relate to well-established definitions of ‘health’ that consider its social and environmental determinants, as in the World Health Organisation (WHO)’s 1948 definition of health as ‘a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity’.⁵⁷

The consequences of datafication, however, mean that such definitions quickly expand to refer to a huge variety of data from which health can be inferred, meaning these practical definitions are no longer meaningful for health data governance and management.

Beyond practical definitions of health data, legal definitions ask: out of all the potential data that relates to health, what data constitutes the legal category of health data, and therefore ought to be subject to special or unique legal provisions?

Under the General Data Protection Regulation (GDPR) in the EU, the specification of ‘special category’ concerns data that is sensitive because its exposure could significantly impact the rights and freedoms of the data subject.⁵⁸ Due to that sensitivity, special category data is data that should be given specified and unique protection under law. Similarly, in US law the Health Insurance Portability and Accountability Act (HIPAA) considers ‘any health information which can be tied to an individual as “protected health information”’.⁵⁹ Across other legal frameworks and jurisdictions, specific conditions exist that must be satisfied to process data about health legally.

57 World Health Organisation. Constitution. Available at: <https://www.who.int/about/who-we-are/constitution> (Accessed: 28 July 2020).

58 Information Commissioner’s Office (2020) ‘Special category data’. ICO. Available at: <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/lawful-basis-for-processing/special-category-data/> (Accessed: 28 July 2020).

59 ‘What is Considered Protected Health Information Under HIPAA?’ (2018) *HIPAA Journal*, 2 April. Available at: <https://www.hipaajournal.com/what-is-considered-protected-health-information-under-hipaa/> (Accessed: 28 July 2020).

However, the datafication of health complicates legal definitions. Laws like the GDPR are designed as principles-based, so that they incorporate a broad range of data under the definition of health data. The GDPR offers non-exhaustive guidance as to what the ‘special categories of data’ might be, such as ‘genetic data’, ‘biological samples’, or ‘medical history’ and refers to traditional clinical actors and settings such as physicians or other health professionals, hospitals and medical devices.⁶⁰ However, as datafication expands the bounds of what can count as health data, such special categories are rendered less meaningful.

This creates a governance gap between data that can be used to make inferences about health and data that is legally considered health data. Increasing datafication means that many uses of health data are at risk of falling through this gap.

This is evident in the USA, where wearables companies are unlikely to be covered under the HIPAA, as they engage directly with customers and do not fulfil the requirement of ‘providing healthcare services’.⁶¹ Because of this, wearables companies are also unlikely implicated under the ‘secondary use’ of health data, where data is sold to advertisers.

In the USA, wearables companies are instead covered by the Federal Trade Commission, and are expected to self-regulate, following guidelines. A legal review of m-health apps found that most of them take private data without seeking explicit permission, and that as a rule they do not disclose that this data will be shared with advertisers.⁶² A 2019 Senate privacy bill proposed to extend the HIPAA to consider more non-HIPAA companies that collect health data.⁶³ However, it only applies to companies with a substantial purpose of collecting personal health data. This means that companies like Facebook and Google, which do not have a stated purpose to collect health data but can infer someone’s health status from their data, would be exempt.

Datafication redefines what counts as sensitive health data, and undermines the practical and legal usefulness of concepts like ‘special category’ or ‘uniquely sensitive health data’.⁶⁴

Rethinking meaningful conceptual, practical and legal definitions of health data means considering the two crucial aspects of the datafication of health discussed earlier:

- **The ability to infer individuals’ health status from data** has increased the sensitivity of all personal data, by changing who can access data about health and what they can know about people’s health and lives.

60 ‘Recital 35 - Health Data’, *General Data Protection Regulation (GDPR)*. Available at: <https://gdpr-info.eu/recitals/no-35/> (Accessed: 28 July 2020).

61 Tomimbang, G. (2018) ‘Wearables: Where do they fall within the regulatory landscape?’ *The Privacy Advisor, IAPP*. Available at: <https://iapp.org/news/a/wearables-where-do-they-fall-within-the-regulatory-landscape/> (Accessed: 28 July 2020).

62 Dehling, T. et al. (2015) ‘Exploring the Far Side of Mobile Health: Information Security and Privacy of Mobile Health Apps on iOS and Android’, *JMIR mHealth and uHealth*, 3(1), p. e8. doi: 10.2196/mhealth.3672.

63 Richardson, T. (2019) ‘Senate Privacy Bill Would Expand HIPAA, Restrict Health Apps’, *Bloomberg Law*. Available at: <https://news.bloomberglaw.com/privacy-and-data-security/lawmakers-float-federal-wearable-devices-privacy-legislation> (Accessed: 6 October 2020).

64 Sobhani, M, Saxon, L (2019) ‘All Our Data is Health Data’, *Medium*. Available at: <https://medium.com/@uscdbc/all-our-data-is-health-data-57d3cf0f336d> (Accessed: 6 October 2020); Warzel, C (2019)

- **The ability to repurpose data about health** has increased the potency of health data and broadened the ways in which data about health can be processed and used for other applications.

However, by broadening the scope of what counts as data related to our health, datafication presents a dilemma:

1. **If practical and legal definitions of health data become broad enough to encompass all potential inferences about health, this may further blur the boundaries between health and other data, and mean that almost all data would have to be regulated as health data.**
2. **If definitions of health data are not expanded, individuals may become vulnerable to having sensitive aspects of their health status available to be inferred and used by different actors beyond the clinical context, without sufficient legal or practical protections.**

Addressing this dilemma requires acknowledging and responding to the consequences of the datafication of health, and reflecting on health data in relation to networked digital technologies used to infer and repurpose it: it requires rethinking what health data is.

4.2 What is health data now?

The datafication of health is changing conceptual, practical and legal definitions of health data, meaning it is now:

1. **Ubiquitous** – Data about health can exist and be generated across any aspect of life that can be made digital.

2. **Comprehensive** – Data about health is increasingly diverse, facilitating a shift in focus from understanding and treating disease to maintaining wellbeing and assisting disease prevention.

3. **Personalised** – Data about health increasingly relates to an individual's health rather than aggregate population-level figures.

4. **Measurement-based** – Quantifiable data about health can be measured, compared, aggregated and used to make profiles, predictions and scores.

1. Ubiquitous

Data about health is now ubiquitous, generated as a by-product of everyday human activity as people's lives have become mediated by online digital services and networked digital devices.⁶⁵

Online digital services – from social networks and forums to shopping sites and fitness blogs – provide insights into people's preferences, behaviour, social structures, personalities, interests and concerns. Digital devices that gather data about health are embedded in people's digital lives, from m-health devices like smartphone apps and wearables, to ambient technologies such as smart speakers and voice assistants, as well as the sensing infrastructure of smart cities. All these networked devices and platforms provide data about users' physiology and psychological characteristics, as well as mapping their movements, sleep patterns and more.

65 Lambiotte, R. and Kosinski, M. (2014) 'Tracking the Digital Footprints of Personality', *Proceedings of the IEEE*, 102(12), pp. 1934–1939. doi: 10.1109/JPROC.2014.2359054.

This offers individuals the opportunity for increased control over their health by providing ‘self-knowledge through numbers’. Self-tracked data has initiated new relationships between data subjects and their bodies, where bodily intuition is being outsourced to disembodied data.⁶⁶ Wearables and health apps play an increasingly central role in the paradigm of health promotion by engaging citizens with knowledge about their health, increasing individual control and facilitating behavioural change.

Ubiquitous networked devices are also driving a shift towards disease prevention, blurring traditional boundaries between clinical care and public health, by providing a scaleable infrastructure for the collection of personal health data. For example, applications such as Apple’s ResearchKit have connected consumers directly to healthcare professionals, allowing medical researchers to conduct studies using the iPhone and Apple Watch.⁶⁷ In 2017, the Apple Heart Study app monitored 400,000 people’s heart rhythms, alerting those at risk of atrial fibrillation.⁶⁸

The use of personal data in situ makes analysis that was previously only available in a clinical setting deliverable on a population-wide scale. However, privacy concerns have been raised around the sharing of personal data between hospitals and tech companies.⁶⁹ Data sharing between University of Chicago Medical School and Google in the USA, and between DeepMind and the NHS in the UK, prompted legal cases, challenging the redrawing of lines as a violation of patient privacy.⁷⁰

The increasing ubiquity of health data also helps to realise a new ideal in healthcare, centred around the involved patient.⁷¹ Widespread data collection facilitates more active participation in health, allowing patients to assess their health and share information with their healthcare provider. This holds the promise of empowering individuals with the ability to know about their health and give them more control and agency over their lives.

However, this ubiquity might also correspond to heightened obligations around maintaining personal health and conforming to narrow conceptions of a healthy lifestyle in order not to be disadvantaged, as well as coming at the cost of allowing technology companies greater access to personal and private data.⁷²

66 Smith, G. J. D. and Vonthehoff, B. (2017) ‘Health by numbers? Exploring the practice and experience of datafied health’, *Health Sociology Review*, 26(1), pp. 6–21. doi: 10.1080/14461242.2016.1196600.

67 Apple. *ResearchKit and CareKit*. Available at: <https://www.apple.com/uk/researchkit/> (Accessed: 6 October 2020).

68 Stanford Medicine. *Apple Heart Study*. Available at: <https://med.stanford.edu/appleheartstudy.html> (Accessed: 6 October 2020).

69 Wakabayashi, D. (2019) ‘Google and the University of Chicago Are Sued Over Data Sharing (Published 2019)’, *The New York Times*, 26 June. Available at: <https://www.nytimes.com/2019/06/26/technology/google-university-chicago-data-sharing-lawsuit.html> (Accessed: 6 October 2020).

70 Information Commissioner’s Office (2020) ‘Royal Free - Google DeepMind trial failed to comply with data protection law’, *ICO*. Available at: <https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2017/07/royal-free-google-deepmind-trial-failed-to-comply-with-data-protection-law/> (Accessed: 12 October 2020); Wakabayashi, D. (2019) ‘Google and the University of Chicago Are Sued Over Data Sharing (Published 2019)’, *The New York Times*, 26 June. Available at: <https://www.nytimes.com/2019/06/26/technology/google-university-chicago-data-sharing-lawsuit.html> (Accessed: 12 October 2020).

71 Appelboom, G. et al. (2014) ‘The quantified patient: a patient participatory culture’, *Current Medical Research and Opinion*, 30(12), pp. 2585–2587. doi: 10.1185/03007995.2014.954032.

72 Erikainen, S. and Chan, S. (2019) ‘Contested futures: envisioning “Personalized,” “Stratified,” and “Precision” medicine’, *New Genetics and Society*, 38(3), pp. 308–330. doi: 10.1080/14636778.2019.1637720.

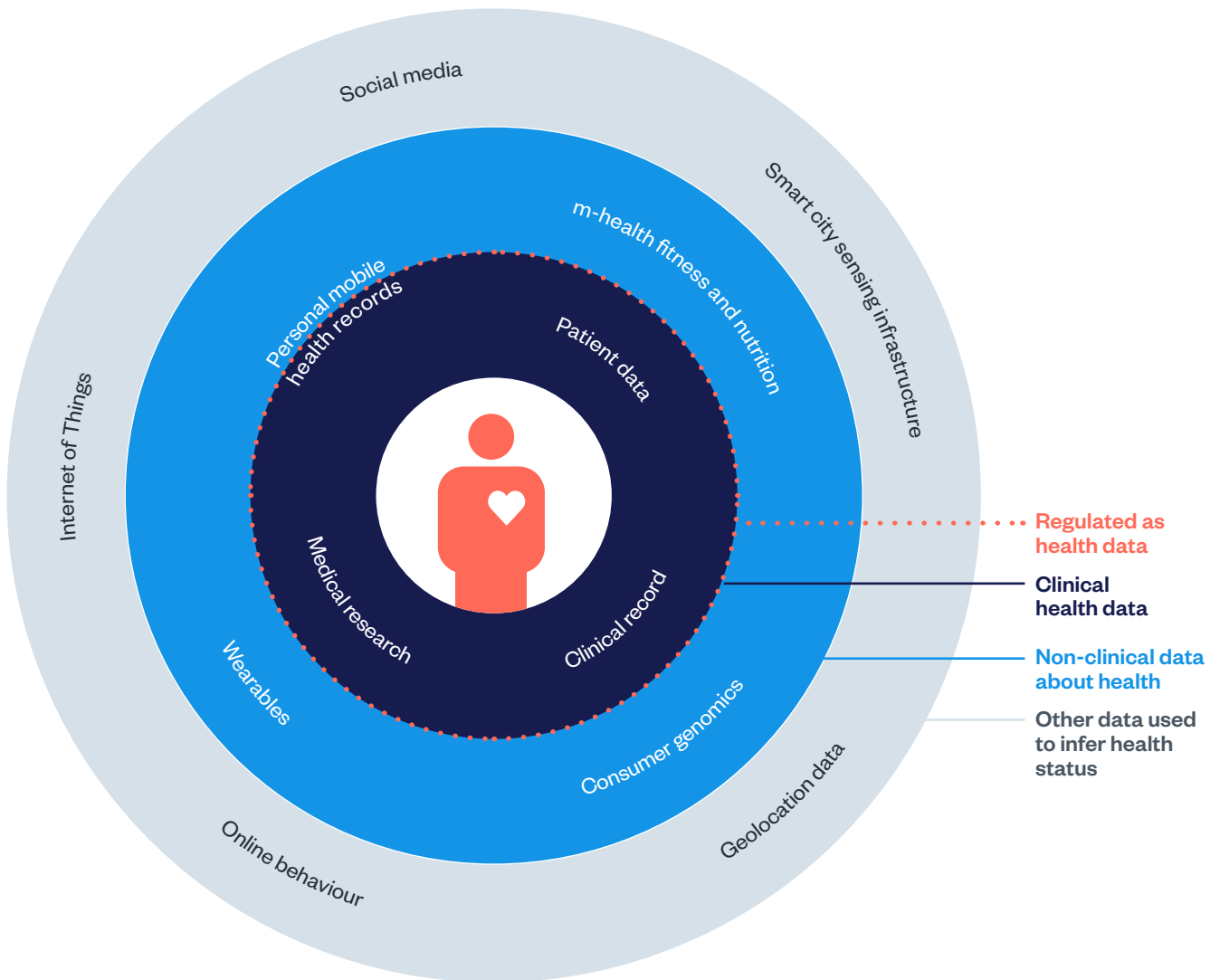


Figure 3
Ubiquitous data collection expands the boundaries of what can count as health data and from where it can be gathered

2. Comprehensive

The datafication of health broadens the range of human activity that can be understood in relation to health. This supports a shift from healthcare that understands and treats disease to comprehensive health monitoring and disease prevention.

The WHO's 1948 definition of health invited nations to 'expand their conceptual framework of their health systems beyond the physical boundaries set by the physical condition of individuals and their diseases.'⁷³ And the modern **Meikirch model of health** describes health as the '*dynamic state of wellbeing* emergent from conducive interactions between individuals' potentials, life's demands, and social and environmental determinants'.⁷⁴

The datafication of health makes advances towards these broad ideals of health: moment-by-moment quantification through ubiquitous devices that gather data means that this comprehensive conceptual understanding of health can be realised.

M-health start-ups are making use of increasingly comprehensive data about health. In 2018, a joint venture emerged between Longevity United Inc's 'healthy living platform' and the Eterly ecosystem, which combines wearables that monitor steps, sleep and heartrate with information about users' eating habits, vitamin intake and mood to determine a 'longevity score'.⁷⁵ The resulting app integrates a machine learning-based 'intelligent chatbot' that functions as a virtual personal trainer and diagnostics tool, incorporating gamification, where users 'mine' tokens by achieving the health goals set by the chatbot.

This new marketplace offers products and services that are integrated with Eterly's recommendations, and is ultimately aimed at establishing cooperation with insurers, drug manufacturers and scientists. Developers of the app state that users will gain from 'personalised health, lifestyle and fitness advice from world leading experts in the field of gerontology, bio-pharma, personalised medicine [...] tailored to them using the latest AI and machine learning'.

This approach to monitoring and maintaining people's wellbeing, while generating a market and advertising platform, is entirely novel and made possible by the increasingly comprehensive nature of data about health.

73 Jadad, A. R. and O'Grady, L. (2008) 'How should health be defined?', *BMJ*, 337. doi: 10.1136/bmj.a2900.

74 Bircher, J., and Hahn, E.G., (2017) 'Meikirch model: A new concept of health based on science'. *Meikirch Model* Available at: <https://meikirch-modell.ch/en/> (Accessed: 6 October 2020).

75 Longevity United (2018) 'AI & Blockchain for Healthy Longevity: Longevity United Launches Joint Venture With Eterly'. *Medium*. Available at: <https://medium.com/@LongevityUnited/longevity-united-and-eterly-are-getting-ready-for-a-new-fundraising-round-586a4746badd> (Accessed: 6 October 2020).

However, this example highlights the limitations of datafied health. The Meikirch model emphasises that health is a result of different interactions within a ‘complex adaptive system’. In other words, datafication allows for the analysis of an incredibly comprehensive range of phenomena related to health, but the process is unavoidably reductive and limited, because health cannot be fully understood solely through quantified metrics. When activity is datafied, people are ‘sliced and diced into decontextualised parts’ and abstracted from their real-world environment.⁷⁶

The reductive nature of seeing health in terms of decontextualised bits of data emphasises that data-driven methods cannot replace traditional patient-doctor relationships that can qualitatively understand people’s embodied experiences of health.

Future approaches to healthcare will require a nuanced approach to the use of data-driven technology, informed by diverse patient perspectives, that supports clinicians and adds value to patient-doctor interactions.

3. Personalised

Thanks to its ubiquity and comprehensiveness, data about health increasingly relates to a *specific individual’s* health, even on a population-wide scale.

Previously, unless an individual underwent tests within a clinical setting, data about health did not relate specifically to *their* health. Rather, data about health would more likely relate to individuals as part of a broader group characterised by statistical similarity, such as having similar physical characteristics or medical risk factors.

The personalised nature of the data generated via the datafication of health plays an instrumental role in the shift away from a ‘one size fits all’ approach to medicine (where a patient is given the medicine that is most likely to work because it often works for people like them), towards personalised medicine that allows for treatments tailored to the individual.

But the concept of personalised medicine is not new. Routine checks of a patient’s blood type before a blood transfusion and consulting family medical history when diagnosing hereditary diseases are both examples of developments in the personalisation of medicine.⁷⁷ With ubiquitous and comprehensive data about health, however, the potential now exists to combine a broad variety of data to identify factors that could help to tailor medical treatments to each patient’s unique biology and environment. A wearable ECG monitor could indicate that a person is at an increased risk of heart disease or stroke: information that they can share with their doctor, providing a more precise and tailored evidence base to inform their treatment.⁷⁸

76 Neff, G. and Nafus, D. (2016) *Self-tracking*. Cambridge, Massachusetts: The MIT Press (The MIT Press essential knowledge series).

77 Ng, M. M. (2015) *The past and future of personalized medicine*, *The Jackson Laboratory*. Available at: <https://www.jax.org/news-and-insights/jax-blog/2015/july/past-and-future-of-personalized-medicine> (Accessed: 6 October 2020).

78 See for example: Withings, *Move ECG – ECG Monitor & Activity Watch*. Available at: <https://www.withings.com/ca/en/move-ecg> (Accessed: 6 October 2020).

Increasingly, personalised approaches rely on seeing the patient alongside their data. P4 medicine aims to associate each patient with a virtual cloud of billions of data points by leveraging information technologies.⁷⁹ Currently, personalised approaches to medicine focus on ‘deep phenotyping’ – analysing an individual’s genetic and biomedical markers and phenotypic abnormalities. However, these markers cannot consider someone’s activity level, stress, social life, dietary habits or living environment.

Datafication offers solutions by combining data from people’s biological phenotype with their digital phenotype, a process called ‘**deep digital phenotyping**’.⁸⁰ This approach could identify unique linkages between diseases that have previously been hidden.

However, digital phenotypes further enable private actors to repurpose health data to make personalised predictions about people. Aspects of a person’s life that haven’t previously been related to health are becoming not just health related but related to fundamental aspects of a person’s biology or genetic make-up. Those biological and genetic factors are legible to those with access to the data.

These practices are not on the horizon but exist already in applications today. Credit scoring has historically been based on a calculation of known financial factors such as payment history and debt, but fintech companies such as Zest AI and Kreditech now assess creditworthiness by analysing users’ digital phenotype, inferring psychological characteristics from social media, contacts, browsing history and geolocation, and repurposing these as an indicator of a potential borrower’s robustness.⁸¹

Though the opportunities to personalise medical care are a positive consequence of datafication, the possibility of repurposing this data raises important questions of individual privacy. Proper scrutiny and protection must be created to set out the contexts in which it is appropriate to process or use data about a person’s health, especially in cases where access to goods and services such as credit might be denied to someone on the basis of their health status.

4. Measurement-based

Datafication redefines and expands the role of health within a broader ‘metric culture’ where numbers and algorithms play an increasingly central role in the governance of people’s lives.⁸² Increasingly comprehensive and personalised understandings of health and ubiquitous health data practices create an unprecedented capacity to measure people and then control, monitor and modulate them, becoming a ‘politics of life itself’.⁸³

79 Vogt, H., Hofmann, B., & Getz, L. (2016). ‘The new holism: P4 systems medicine and the medicalization of health and life itself.’ *Medicine, health care, and philosophy*, 19(2), 307–323. <https://doi.org/10.1007/s11019-016-9683-8>

80 Fagherazzi, G. (2020) ‘Deep Digital Phenotyping and Digital Twins for Precision Health: Time to Dig Deeper’, *Journal of Medical Internet Research*, 22(3). doi: 10.2196/16770.

81 Zest AI. ‘Transform Your Underwriting with AI & Machine Learning’, Available at: <https://zest.ai/> (Accessed: 6 October 2020); Kreditech. Available at: <https://www.kreditech.com/> (Accessed: 6 October 2020).

82 Ajana, B. (2018) ‘Introduction: Metric Culture and the Over-examined Life’, in Ajana, B. (ed.) *Metric Culture*. Emerald Publishing Limited, pp. 1–9. doi: 10.1108/978-1-78743-289-520181001.

83 Rose, N. S. (2007) *Politics of life itself: biomedicine, power, and subjectivity in the twenty-first century*. Princeton: Princeton University Press (Information series).

This can be understood in terms of concepts of a risk society.⁸⁴ Economic, environmental, security and health risks all dominate modern public discourse. This discourse associates lack of knowledge with a failure to anticipate risk and proposes that greater knowledge enables the mitigation and control of risk. In accordance with the motivations of a risk society, different actors increasingly mobilise data to understand, mitigate and predict risks.

In this context, data about health is not only a personal and medical asset but also a financial asset, where knowing about people's health can mitigate risk. For example, corporate wellness schemes that leverage wearable technologies to measure employee health exemplify 'knowing about health' as a way of mitigating economic risk.⁸⁵ In this case the mitigation of risk is achieved for the employer, by shifting the risk on to the individual, who might lose out on certain health or financial benefits if their statistics don't meet certain criteria.

The personalised nature of data about health also leads to a decollectivisation of risk, where granular, individualised risk profiles can be created from aspects of a person's digital phenotype.

In 2019 John Hancock, one of the largest North American life insurers, announced its intention to stop underwriting traditional life insurance policies, instead selling interactive policies that track fitness and health through wearable devices and smartphones.⁸⁶ Such interactive policies increase scrutiny on people, potentially disempowering them, as obligations to maintain their health manifest through increased surveillance and control.

The expansion of the clinical gaze via data has the potential to erode individual and collective freedoms, shaping lives according to measurable risk factors, subordinating other valid goals to the specified requirements of a healthy life. This represents a seismic shift: whereas 'medicine used to claim authority over the cracks and interruptions in life; now it claims authority over all of life'.⁸⁷

Digital phenotyping and the measurement of health through data from individual activity, promotes an implicit assumption that health ought to be managed by the individual. This conception of health ignores the structural determinants of health and entrenches health inequalities because they cannot be as easily measured. In a risk society, the incentives do not exist to design systems and use data in a way that accounts for these factors, which places responsibility for risk back on to authorities rather than distributing it among individuals.

-
- 84 Beck, U. (1992) *Risk society: towards a new modernity*. London ; Newbury Park, Calif: Sage Publications (Theory, culture & society).
- 85 BasuMalik, C. (2019) 'Why Wearable Technologies Will Drive Workplace Wellness in 2019'. *HR Technologist*. Available at: <https://www.hrtechnologist.com/articles/workplace-wellness/how-wearable-technologies-will-drive-workplace-wellness-in-2019/> (Accessed: 6 October 2020).
- 86 Chen, A. (2018) 'What happens when life insurance companies track fitness data?', *The Verge*. Available at: <https://www.theverge.com/2018/9/26/17905390/john-hancock-life-insurance-fitness-tracker-wearables-science-health> (Accessed: 6 October 2020).
- 87 Downing, R. (2011) *Biohealth: beyond medicalization : imposing health*. Eugene, Or.: Pickwick Publications. p.15.

By ignoring structural inequalities and attending to an idealised image of individuals as a free, self-optimising agent, the wellness ideal, driven by the datafication of health, embodies the vision of the ideal neo-liberal subject.⁸⁸ Wellness co-opts narratives of ‘empowerment’, while encoding a rampant individualism – that each person alone is responsible for their health.⁸⁹ For example, Fitbit’s motto is ‘To empower and inspire you to live a healthier, more active life.’⁹⁰

Such sentiments assume that a person *can* be empowered or inspired into health, while implicitly ignoring the fact that someone with a chronic illness or someone who works multiple jobs to pay their bills, faces different challenges to be ‘inspired’ into a healthier, more active life, no matter how much data is gathered about them, and used to measure their health and score their risk levels.

88 Ajana, B. (2017) ‘Digital health and the biopolitics of the Quantified Self’, *Digital Health*, 3, p. 2055207616689509. doi: 10.1177/2055207616689509.

89 McBain, S. (2020) ‘The dark side of the wellness industry’, *The New Statesman*. Available at: <https://www.newstatesman.com/politics/health/2020/06/dark-side-wellness-industry> (Accessed: 6 October 2020).

90 Fitbit. *About Fitbit*. Available at: <https://www.fitbit.com/global/us/about-us> (Accessed: 6 October 2020).

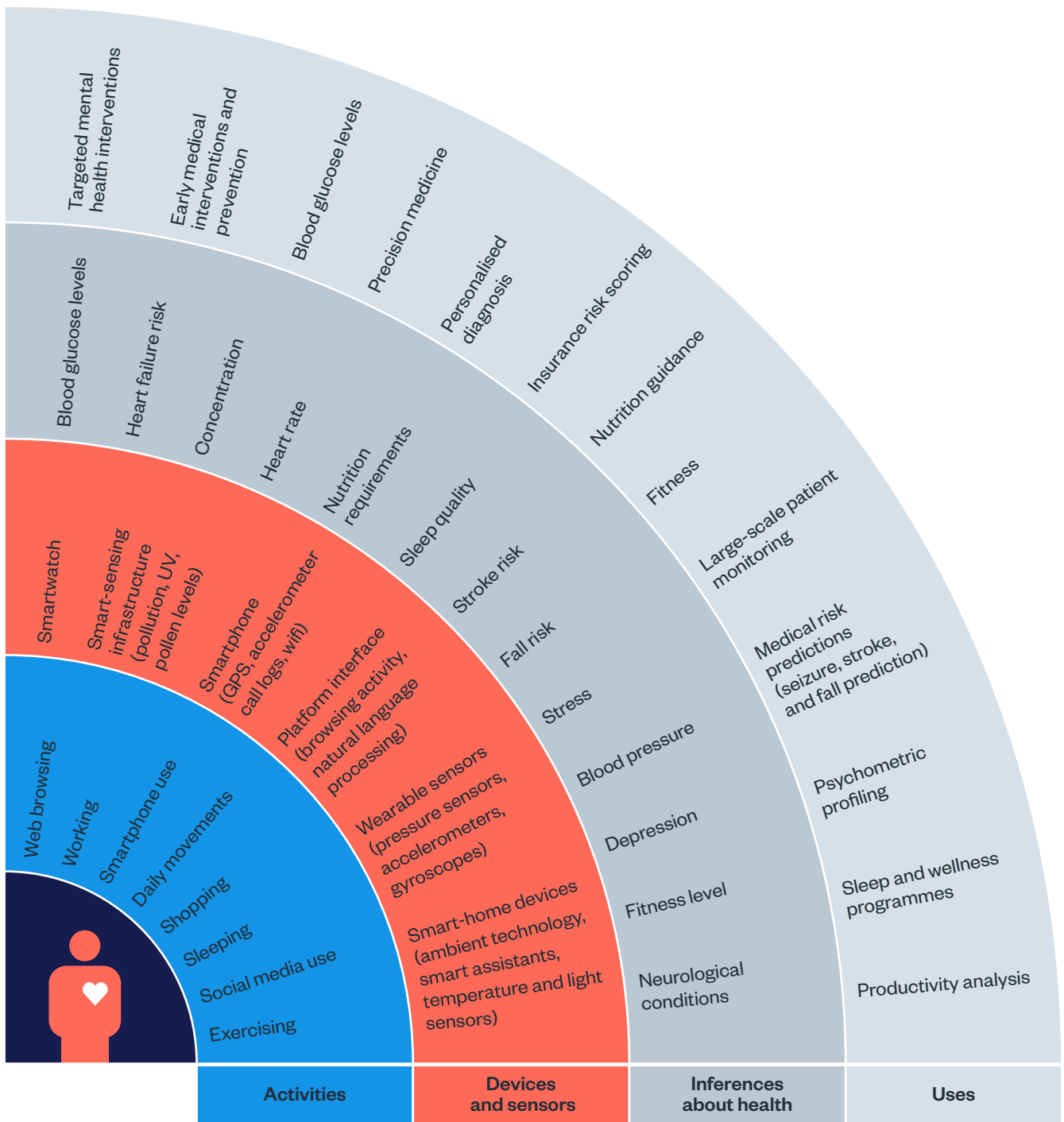


Figure 4
 The datafication of health by ubiquitous technologies which monitor human activity makes it possible to make inferences about health, which can be used for a range of purposes. This builds a comprehensive picture of health.

5. The Internet of Health

Health data is no longer confined to the domains of hospitals and medical laboratories, clinical trials and public health organisations. Just as vast networks of connected smart devices have created an Internet of Things, through ubiquitous devices that gather comprehensive, personalised and measurement-based data about health, datafication has created an 'Internet of Health'.

Within the traditional boundaries of health, individuals became patients in regulated clinical settings where health data might be collected and processed from consultations, blood tests, clinical trials or scans. In the Internet of Health, however, everyone's health is subject to continuous monitoring, research, diagnosis and prediction through data. Everybody is a patient whether they like it or not.

In this Internet of Health, data about health can be continuously gathered, inferred, processed and used by actors with the appropriate digital tools to access and analyse health data. Health is made legible to those who have access to the data, which includes an app user who logs their exercise and tracks their progress, as well as the company that provides the app.

However, if everybody is a patient in the Internet of Health, governments and societies must seriously consider the 'duty of care' that corporations have towards individuals whose health becomes legible through the processing of their data. If corporations are profiting from data about consumers' health, do they have a correlative duty to inform data subjects of potential serious health risks? Does a supermarket chain have a duty to inform an individual if their shopping data flags that they have a high cancer risk?

Within traditional clinical boundaries, individuals relinquish a great deal of control and defer judgement to actors with the ability to know about health: to doctors, nurses, health organisations and researchers with specialised skills and expert judgement. This decision is made in the knowledge that strict regulatory oversight and codes of conduct protect patients against being disempowered or mistreated. In the Internet of Health, however, individuals relinquish control over their data and health, but without the ecosystem of regulation, protection, oversight and care that exists within clinical settings.

Increased legibility of people's health in the Internet of Health not only blurs regulatory boundaries, but also dissolves the intuitive, contextual boundaries structuring individuals' privacy expectations around health data. In the Internet of Health, individuals are vulnerable to disempowerment, where knowledge and power asymmetries leave them unable to maintain control over both their health and the data about it.

The ability to repurpose data about health also causes concerns for individual agency, through subordinating individuals' healthcare interests to corporate productivity, and to lucrative prediction markets. Individuals are nudged by increasingly personalised, automated systems that mobilise medical authority. Data about their health is linked and connected, which creates **networked health identities** that are categorised according to risk. These networked health identities become a source of discrimination and disadvantage, where data about health is repurposed to deny certain people goods and services.

The datafication of health, and the Internet of Health it creates, requires an urgent rethink of the legal and practical definitions of health data. It offers considerable advances in health and care, but only if the challenges, concerns and questions it raises are fully addressed, and individuals are not harmed or disempowered by the proliferation of health data collected about them.

Many of these examples we refer to are from the EU, the UK and the USA, to situate the analysis within the context of GDPR and acknowledge the influence of US-based technology companies that drive datafication. But the consequences of datafication manifest wherever it occurs, and attention to its impacts across cultures and national settings remains in need of research.

We conclude this report by outlining a series of questions that act as a foundation to address an overarching, urgent challenge raised by the datafication of health: **what does it mean to be a patient in the Internet of Health?**

These questions build on the concepts synthesised in this report and lay a foundation for future work by researchers, policymakers, technology developers and health professionals to help ensure data about health works for people and society:

Empowering individuals and supporting agency

1. How can technologies and tools involved in the datafication of health be designed in ways that promote individual agency and wellbeing rather than market-driven predictions or control?
2. How should existing legal frameworks that ensure people's data rights be amended to account for the expansion of what counts as health data, especially where health data is afforded special or unique status?
3. What models of governance in an Internet of Health will give people greater influence over their own health, and over developments and treatments for conditions they have?

Ensuring equity of health opportunities and outcomes

4. What practical safeguards and legal frameworks are needed to ensure that societal or political values in the design of health data systems do not embed bias and discrimination in the processes that shape how people are made legible through health data?
5. How can the benefits of datafication be leveraged to minimise health inequalities and improve everyone's access to good quality healthcare, without undermining their agency?
6. How do different communities (of health, culture, gender, ethnicity) exist and interact within the Internet of Health, and are there equitable opportunities to participate?

Bolstering solidarity and societal wellbeing

7. How can the interests of individuals, organisations, companies, corporations and governments be balanced in a way that promotes empowerment and knowledge rather than creating burdens of self-care and disempowerment?
8. What incentives for people, organisations and governments will meaningfully promote the benefits of datafication, and reduce the burden of its costs, to enhance and improve health for everyone?
9. How does datafication affect relationships and influence power asymmetries between people and organisations who gather and use data about health?

Developing inclusive and participatory health data practices

10. How should data-driven systems be designed to ensure the experiences of those who fall outside of the category of a technology's 'typical user' are not excluded?
11. How does the datafication of health minimise or create opportunities for individual expressions and experiences of health identity? What opportunities are there in the Internet of Health to increase the ability for people to participate in and influence the structures that govern health and data?
12. What opportunities are there in the Internet of Health to increase the ability for people to participate in and influence the structures that govern health and data?

Glossary

Key terms	Definitions
Data about health is now: comprehensive	Data about health is increasingly broad and diverse, facilitating a shift in focus from understanding and treating disease to maintaining wellbeing and disease prevention.
Data about health is now: measurement-based	Quantifiable data about health can be measured, compared, aggregated and used to make predictions and scores.
Data about health is now: personalised	Data about health increasingly relates to an individual's health rather than aggregate population-level figures.
Data about health is now: ubiquitous	Data about health can now exist and be generated across any aspect of life that can be made digital.
Datafication	The recording of people's everyday life in quantified data.
Datafication of health	The process where individuals' activity, behaviour and experiences are recorded in quantified data and made analysable in terms of health.
Digital phenotyping	The moment by moment quantification of people's phenotypes using data from personal digital devices.'
Deep digital phenotyping	The combination of deep phenotyping, defined by Robinson in 2012, as the 'precise and comprehensive analysis of phenotypic abnormalities in which the individual components of the phenotype are observed and described' with digital phenotyping.
Hypernudge	A Hypernudge, as defined by Karen Yeung in 2016, is a technique which combines big data analytics with personalisation to produce highly persuasive and personalised influences on a person's behaviour.
Inferring health	Making inferences about a person's health from data which is not classified as 'health data' and which has been collected outside of a traditional healthcare context.
The Internet of Health	A new networked health domain, created by the datafication of health, which has made individuals' health status legible to actors with access to their data.
Legibility	Legibility, as defined by Linnet Taylor, is 'the way in which citizens become visible, or, legible, to authorities (public or private) through data collection and analysis.
Meikirch model of health	The Meikirch model describes health as the 'dynamic state of wellbeing emergent from conducive interactions between individuals' potentials, life's demands, and social and environmental determinants.'
Networked Health Identity	The combination of inferences about a person's health, generated from the data they produce across networked platforms and devices, which can be used to categorise them according to risk or other factors.
P4 medicine	A vision of medicine that is predictive, personalised, preventative and participatory.
Patient activation	The knowledge, skills and confidence a person has in managing their own health and care.
Phenotype	The term used in genetics for the observable characteristics of an individual that result from the interaction of their genotype (the collection of genes an individual has) with their environment.
Repurposing health	Using data about health for purposes not directly related to healthcare.

Acknowledgements

We would like to thank the following colleagues for taking time to review a draft of this paper and offering their expertise and feedback:

- Dr Rachael Kent, King's College London
- Dr Jared Keller, Dr Mahlet Zimeta and colleagues, Open Data Institute
- Nisha Shah, Centre for Health, Law and Emerging Technologies (HeLEX), University of Oxford
- Sobia Raza, Health Foundation
- Tariq Khokhar, Wellcome Trust

About the Ada Lovelace Institute

The Ada Lovelace Institute was established by the Nuffield Foundation in early 2018, in collaboration with the Alan Turing Institute, the Royal Society, the British Academy, the Royal Statistical Society, the Wellcome Trust, Luminata, techUK and the Nuffield Council on Bioethics.

The mission of the Ada Lovelace Institute is to ensure that data and AI work for people and society. We believe that a world where data and AI work for people and society is a world in which the opportunities, benefits and privileges generated by data and AI are justly and equitably distributed and experienced.

We recognise the power asymmetries that exist in ethical and legal debates around the development of data-driven technologies, and will represent people in those conversations. We focus not on the types of technologies we want to build, but on the types of societies we want to build.

Ada Lovelace Institute

Part of the Nuffield Foundation
Registered charity 206601

ISBN: 978-1-8382567-0-8

Through research, policy and practice, we aim to ensure that the transformative power of data and AI is used and harnessed in ways that maximise social wellbeing and put technology at the service of humanity.

We are funded by the Nuffield Foundation, an independent charitable trust with a mission to advance social well-being. The Foundation funds research that informs social policy, primarily in education, welfare and justice. It also provides opportunities for young people to develop skills and confidence in STEM and research. In addition to the Ada Lovelace Institute, the Foundation is also the founder and co-funder of the Nuffield Council on Bioethics and the Nuffield Family Justice Observatory.

adalovelaceinstitute.org
[@AdaLovelaceInst](https://twitter.com/AdaLovelaceInst)
hello@adalovelaceinstitute.org