From Spurling, N. and Kuijer, L. (2016) Everyday Futures. Institute for Social Futures: Lancaster
http://wp.lancs.ac.uk/everydayfutures/essay-collection/

"By their [data] you will know them"[i] – Historical reflections on capturing patterns in everyday life

Derek Gatherer (Health and Medicine, Lancaster University, UK) d.gatherer@lancaster.ac.uk
Lenneke Kuijer (Industrial Design, TU Eindhoven, Netherlands) s.c.kuijer@tue.ac.uk
Ida Nilstad Pettersen (Product Design, NTNU, Norway) ida.nilstad.pettersen@ntnu.no

Abstract
With large quantities of digital data collected on our everyday lives, concerns arise as to how this data may affect these very lives. To derive relevant research questions concerning Everyday Futures, our essay reflects on the use of digital data in everyday decision-making. We do so by comparing historic and contemporary examples of health related data-action loops on three different scales: the body, the home and the city. We conclude that while the use of data to inform sensitive decisions is not new, digitization gives rise to a number of important research themes, including tensions between developers and users, theory and opportunity, sensors and senses, and norms and diversity, 'expert' and actor, and that what is (thought to be) measured versus what is not. Moreover, we illustrate how our multiscale, historic, multidisciplinary reflection forms a potential method for everyday futures research.

Introduction
The Jawbone UP® system is one of many wearable devices[ii] that use sensors to collect a range of personal data that are communicated in a related app, where they are accompanied by group averages and health recommendations[iii]. UP® is argued to 'help people live better by providing personalized insight into how they sleep, move and eat'[iv]. But the system does more than merely offering advice. In their alliance with the NEST® thermostat, the UP® service uses the ability to operate home heating systems to take action towards 'making [these better lives] reality'[v]. When the wristband identifies 'sleep mode', the thermostat will automatically 'kick down' the bedroom temperature to 'your ideal temperature', which, according to UP® instructions lies between 18.3-22.2C. Besides this recommended temperature range, the application seems to include a certain enforcement of 'healthy' sleeping temperatures; from an independent review[vi] we learn that the system does not allow a sleep-setting below 15.5C.

What we see in this example is that digital data is used not only to give people insight into quantified indicators of their own actions, but also that the data enters a wider loop where it is aggregated into databases and translated into group averages. Moreover, a system of digital sensors and actuators uses a combination of situated measurements, pre-set preferences and 'expert' recommendations to autonomously take action in people's homes.

In this paper we reflect on the possible effects of digitization for such 'data-action loops' – from collection through to action – in shaping (future) everyday life. We do so by comparing historic and contemporary examples on the use of data for informing health related decisions on three different scales: (1) the body, comparing phrenology, believed in the early 19th century to predict character, with human genome data in which considerable hope is invested for the prediction of future health, (2) the home, comparing practices of domestic heating in the 1920s with the latest smart thermostats that aim to automatically provide a healthy indoor climate; and (3) the city, comparing the 1854 mapping of the London cholera outbreak with a service collecting data on urban mobility. These comparisons are then used to derive questions for everyday futures research.

Data of the Body, the Body of Data, the Body as Data
Long before the advent of electronic medical data collection and storage[vii], the recording and cataloguing of data on the human body was an abiding preoccupation in many societies. The "Bills of Mortality" produced

Please use the following format if you wish to cite this essay:
Gatherer, D., Kuijer, L. and Nilstad-Pettersen, I. (2016) "By Their Data You Will Know Them: Historical Reflections on Capturing Patterns in Everyday Life". In: Spurling N. and Kuijer, L. (Eds) Everyday Futures. Lancaster: Institute for Social Futures. http://wp.lancs.ac.uk/everydayfutures/essay-collection/, pp.59-64

by mediaeval European cities served as both a primitive census of the city's inhabitants and a guide to why mediaeval cities almost always ran a population deficit, needing to import citizens to maintain numbers. As mediaeval society evolved into the early modern state, the data collected on the citizenry moved from basic head-counting to details of the individual. Conscripted armies needed to know how tall, short or malnourished their raw material was, if only for the simple reason that uniforms and equipment of the correct size had to be procured. Poorhouses needed to know how old and infirm their occupants were, if only to assess their ability to cover their maintenance costs through manual work. However, the data soon began to tell its own stories. Italian conscript records for example, provided a striking documentation of improvements in health among young 19th century males[viii]. A decrease in consanguineous marriages across Europe revealed the influence of railways and bicycles on rural life[ix].

Once it was realised that personal data documented change, its potential to implement change became obvious. The phrenology movement, which had its heyday from around 1790 to 1840, didn't just count heads, but studied their shape – and that shape became the basis for theories of personality, criminality and human potential that sought to proactively identify how people might behave in the future. As such, human data had become predictive. At the same era, biochemists began to treat blood and other bodily fluids as objects for laboratory investigation and physiologists began to document the electrical signals passing along nerves and into muscles. The "allure of numbers" began to cast its shadow over medicine[x].

As the 20th century dawned, some of the components of this nascent data science faded away. Phrenology had failed to provide the accurate forecasts it promised, but into the gap created by phrenology's departure from the scientific mainstream came genetics, molecular biology, and, by the mid2000s, "deep" biology or "omics" – a high speed, high capacity data crunching science that sucks terabytes of data from its subjects: their genomes, how those genes are regulated, the genomes of all the microbes within them, how it all interacts, and what this means.

"Yes we can" was one of the political catchphrases of the Obama era in the USA and "because we can"[xi] seems to be a significant part of the philosophy behind deep data collection. For instance, the Precision Medicine Initiative® (PMI) aims to collect genetic and molecular information from enough people to create a picture of how health and disease are mirrored in vast constellations of bio-marker data[xii]. PMI® aims to recruit and follow 1 million volunteers at a cost of approximately $215 million dollars to the US taxpayer in the first year alone[xiii]. Deep biology is not looking for anything in particular; its goal is not specific answers but the big picture: a new, overarching, quantitative definition of what it means to be healthy and how that data documents its change into disease[xiv]. Blood, saliva, urine and faecal samples could be collected every three months for around 100 biochemical tests[xv]. Physiological testing could include sleep patterns, physical activity and heart rate. The subjects' genomes could be sequenced, along with their microbiomes – the genomes of all the bacteria in their guts and in the various body fluid samples. This data could be mined for correlations between states of health and shifting patterns among the bio-markers.

By 2045, the array of omics technologies deep biology employs may have been developed to a point where they can be scaled down to wearable devices. Jawbone UP® and other technologies have already achieved this for many of the physiological data points (sleep, activity, body temperature, pulse; in the future perhaps even blood pressure) required by a deep biology project. All the data could transfer wirelessly to compute clouds trained on the aggregate bio-marker patterns. In such a system, measurable aspects of the workings of the body could be perpetually monitored and analysed, like the engine of a Formula 1 car during a race. Instead of booking an appointment with the doctor, the network might then identify an anomaly, book the appointment and notify you by text before you have even begun to feel unwell.

Data in the Home, the Home of Data, the Home as Data

To safeguard the health and comfort of their residents, homes in temperate or cold climates have always required a form of heating during spells of colder weather. While the solid fuel fire has served this purpose since as long as one million years ago, the past century has seen major changes in home heating technologies. Fundamentally, heating of the home has remained a matter of converting fuel into appropriately heated spaces, but the ways in which decisions, whether deliberate or routinized, about when to add how much fuel to the heating system have changed considerably, as well as ideas of what an appropriately heated space entails. Because there is a time-lag between the acts of adding fuel to the heating system and reaching the desired effect of a warmer room, the decision to add fuel involves a combination of assessing current climate circumstances and comparing them with anticipated near future desired ones. The following sections describe changes in heating systems in more detail while focusing on this decision, and the 'data' used to inform it.

With the shift from coal fires to gas heating, which in Europe took place somewhere between the late 19th century and the 1980s, along with the spreading of gas infrastructures to homes – a first, small, but fundamental shift in decision making took place. In the case of a wood or coal fire, the decision whether to add fuel was made by household members based on a wide range of 'data' in the home that was available to humans, such as situated information (temperature sensations, visual overview of the space, ability to use and interpret speech and non-verbal body language of other household members) and memories of information collected at other moments in time, such as the level of fuel stock, past patterns of household life, temperature preferences, public health advice about indoor climate, experiences of indoor climate in other households, and other snippets of information.

With the introduction of gas fires, the task of adding fuel to the fire was delegated to gas pipes and a simple system that regulated the gas flow based on the setting of a knob. While people still made the decision to set the knob, the setting of the knob, made in a certain moment, now determined whether fuel was added to the fire regardless of other changes in circumstances. With the spreading of thermostatic control, a particular aspect of situated circumstances, namely air temperature re-entered the automated decision to add fuel. However, this also meant a major change in the practice; while the quantifiable and digitally measurable indicator of air temperature had played no or only a minor role before, it has since become prominent in the practice.

When programmable thermostats entered the home around the 1970s, the relation between data, interpretation and action changed again. Programmed thermostats contain anticipated, appropriate temperature levels for different moments in the day and week for the household. When the system is in operation, this data is supplemented with situated temperature measurements, and the occasional overruling of the system by a person that is too warm or cold. In this configuration, sensor data on temperature levels and aggregated patterns of living anticipated by household members, or in the case of default settings, by the system developers take precedence as input for the decision when to add fuel to the heating system.

The next generation of heating systems, 'smart' thermostats, such as the NEST®, expands the variety of situated data collected by sensors, from temperatures to motion and humidity, and store weekly patterns of occupancy – derived from motion sensor data – to make predictions on future occupancy. The assumption that household life shows regular weekly patterns of occupancy, absence and sleep forms the foundation of such systems.

Compared to the solid fuel fire situation, the data that is used to inform the decision whether or not to add fuel to the fire to create an appropriate indoor climate has shifted considerably. From human-sensed holistic circumstances and estimated, near future and needs for heat, to measured temperatures and aggregate occupancy schedules. In the process, what constitutes an appropriate, healthy room climate has also changed. The possibility to continue adding fuel to the fire in the absence of people has given rise to the expectation to always enter a warm room, and with the spreading of thermostatic control, standards of healthy indoor climate control have converged into numerical values. With the Jawbone® UP® technology, another step is taken towards using aggregate, de-contextualised data to 'act' in specific everyday contexts. In particular, ideas of what is an appropriate indoor temperature become further removed from the situated circumstances in the home, and derived more strongly from an aggregate or 'expert' idea of appropriateness.

Data in the City, the City of Data, the City as Data

As the Jawbone® example shows, data collected on the state of bodies and homes are used to inform decision-making, in pursuit of goals about better living. At the city level, visual representations of data have been presented as ways of understanding issues ranging from the spread of diseases to the ways in which people travel. The 1854 London cholera outbreak is frequently used as an example of how visual representations of data can lead to new insight: searching for the source of the outbreak, Dr John Snow plotted the location of the deaths on a map[xvi]. Marking the water pumps in the area on the map, he observed that most of the infected individuals had lived close to a specific water pump. At the time, cholera was thought to be airborne, but aided by his visualisations Snow is thought to have developed the hypothesis that the water was infective, and the pump the source of the outbreak[xvii]. Others do however argue that Snow had a pre-existing theory about cholera being waterborne: Coupled with experience from former outbreaks, his own local knowledge and information he got from talking to people, not the data visualisation alone, his theory contributed to the insight that the pump was the source.

A contemporary example of a data-driven approach to health on a city level is that of Strava Metro®[xviii], a service aggregating anonymised data from individuals who have tracked their walks, runs and bicycle rides by means of apps or wearable GPS devices such as GPS watches. Information about users, their private activities and every first and last kilometre is removed to avoid that it can be associated with specific individuals[xix]. The firm experiments with ways of visualising such data for example, by creating publicly available heat maps[xx] that indicate “the best roads and trails worldwide”[xxi]. Through the service Strava Metro®, data on where people go, how and when – minute by minute, is sold to transportation and city planning departments. This is presented as a tool for “data-driven bicycle and pedestrian planning”, and better data as a “catalyst for change”[xxii]: it may help city planners understand how people use cities, as input for making infrastructural improvements, and to evaluate the effects of their interventions.

However, reflecting on the example in light of our other cases, the service provides a selection of mobility data collected as part of exercise and competition, which means that certain trips and demographic groups may be overrepresented, and others completely absent. This may be problematic, if the use of roads and trails in the sample is interpreted as a vote[xxiii] reflecting the general quality of those roads and trails. Therefore the service does not simply replace alternative quantitative and qualitative tools enabling data collection on mobility. Still, “smart data” such as that provided by Strava Metro® may be used to inform decision-making as a source representing city use, and as a basis for predicting it. The use of “normal” activity patterns deduced from digital data to implement change thus extends to the urban environment.

Conclusions

In this essay we set out to explore what opportunities for everyday futures research a proliferation of digital data collection in everyday life presents. Our quick, multiscale, historic and multidisciplinary reflection on a number of specific, health related examples rendered a range of themes and questions.

First, our analysis of the use of data on different scales helped identify overarching trends related to digitization. Recurring themes were the role of aggregate data in setting a numerical norm, which seems to be increasingly used to inform decisions on all scales, sometimes even made by artefacts. Also, increasing opportunities to generate quantifiable indicators seem to expand the role of data that can be collected with sensors over others, such as the human senses. What if decisions on how to act in unique, everyday situations become increasingly informed by norms determined by averages or 'experts'? What are the effects on daily life of what is thus brought in – or assumed to be, and what is left out?

Second, our historic analysis showed that the use of data in health related decisions is not new. But while data in earlier examples is used to explain historic events, over time, measurable indicators were increasingly used to predict future events. Moreover, on all scales we see situations where abundantly available data is used to *generate* the questions to which the data contains an answer. From this observation, questions arise such as what if these indicators are as predictive as skull measurements are for character? And what place does theory have in these data-intensive attempts at understanding and controlling bodies, homes and cities? Which questions, in such a system, remain unasked?

Finally, our different academic backgrounds, in industrial design and biology, also brought insights on the table, such as analogies between artificial and biological systems. One of the effects of digitization seems to be that it enables us to make our machines more lifelike, maintaining equilibrium or stability, and to generate yet more data-intensive models of biological/medical systems. The city thereby edges closer to the sort of control that is intrinsic to living systems, and which is increasingly part of smart homes. This analogy conjures up questions like: will the future city be able to control its circulation (of traffic and individuals rather than blood) and its respiration (by stopping traffic when air quality deteriorates)? Will human diseases become engineering problems and mechanical breakdowns pathologies? What would this imply for mundane everyday life and what is considered healthy and normal?

Acknowledgements
Part of the historic data used in the domestic space heating case was collected in the context of DEMAND Centre research on 'Adapting infrastructures for a lower carbon society'.

Endnotes

[i] Matthew 7:16. The original quote reads 'By their fruit you will know them'.
[ii] coined in 2015 as one of the fastest growing consumer electronics products by BrandView with an 11 per cent penetration rate in the USA
[iii] https://jawbone.com/blog/up-by-jawbone-insights/
[iv] https://jawbone.com/about
[v] https://jawbone.com/blog/jawbone-up-works-with-nest/
[vi] http://www.telegraaf.nl/digitaal/gadgets/23567492/__Review__Nest__.html
[vii] For a summary of current wearable technology, see http://www.nature.com/nbt/journal/v33/n5/fig_tab/nbt.3222_F1.html
[viii] A'Hearn B, Peracchi F, Vecchi G. (2009) Height and the normal distribution: evidence from Italian military data. *Demography*. 46:1-25.
[ix] Cavalli-Sforza LL, Moroni A, Zei G (2004) *Consanguinity, Inbreeding, and Genetic Drift in Italy*. Princeton University Press.
[x] Winter, JM. (1980) Military Fitness and Civilian Health in Britain during the First World War. *Journal of Contemporary History* 15: 211-244
[xi] https://www.nih.gov/precision-medicine-initiative-cohort-program/infographics declares that "the time is right because: we have a greater understanding of human genes, people are more engaged [..and..] we have the tools [..and..] large databases [..and..] research technologies have improved".
[xii] http://www.sciencemag.org/news/2014/07/google-x-sets-out-define-healthy-human
[xiii] https://www.nih.gov/precision-medicine-initiative-cohort-program
[xiv] http://bmcmedicine.biomedcentral.com/articles/10.1186/s12916-014-0239-6
[xv] https://www.systemsbiology.org/research/100k-wellness-project/
[xvi] Tufte, E.R. (2001) *The Visual Display of Quantitative Information*, 2nd ed. Graphics Press, Cheshire, Connecticut.
[xvii] Brody, H.; Rip, M.R.; Vinten-Johansen, P.; Paneth, N.; Rachman, S. (2000): Map-making and myth-making in Broad Street: the London cholera epidemic, 1854, *The Lancet* 356: 64-68.
[xviii] Strava Metro, http://metro.strava.com/ [accessed 31.08.16]
[xix] Strava Support/Gordon, M. (2016): What is Strava Metro? https://support.strava.com/hc/en-us/articles/216918877-What-is-Strava-Metro- [Accessed 31.08.16] The Guardian/Peter Walker (2016): City planners tap into wealth of cycling

data from Strava tracking app, https://www.theguardian.com/lifeandstyle/2016/may/09/city-planners-cycling-data-strava-tracking-app [Accessed 31.08.16]

[xx] Strava LABS, http://labs.strava.com/heatmap/#6/-120.90000/38.36000/blue/bike [Accessed 31.08.16]

[xxi] http://labs.strava.com/projects/ [Accessed 31.08.16]

[xxii] Strava Metro (n.d.): Data-Driven Bicycle and Pedestrian Planning, http://cdn2.hubspot.net/hubfs/1979139/Strava_Metro___Data-Driven_Planning.pdf Strava Metro (2015): Comprehensive User Guide, Version 2.0 for 2015, http://ubdc.ac.uk/media/1323/stravametro_200_user_guide_withoutpics.pdf

[xxiii] http://metro.strava.com/faq/ [Accessed 08.09.16]