

PERSPECTIVE

Digital footprints: facilitating large-scale environmental psychiatric research in naturalistic settings through data from everyday technologies

N Bidargaddi^{1,2}, P Musiat¹, V-P Makinen², M Ermes¹, G Schrader¹ and J Licinio^{1,2}

Digital footprints, the automatically accumulated by-products of our technology-saturated lives, offer an exciting opportunity for psychiatric research. The commercial sector has already embraced the electronic trails of customers as an enabling tool for guiding consumer behaviour, and analogous efforts are ongoing to monitor and improve the mental health of psychiatric patients. The untargeted collection of digital footprints that may or may not be health orientated comprises a large untapped information resource for epidemiological scale research into psychiatric disorders. Real-time monitoring of mood, sleep and physical and social activity in a substantial portion of the affected population in a naturalistic setting is unprecedented in psychiatry. We propose that digital footprints can provide these measurements from real world setting unobtrusively and in a longitudinal fashion. In this perspective article, we outline the concept of digital footprints and the services and devices that create them, and present examples where digital footprints have been successfully used in research. We then critically discuss the opportunities and fundamental challenges associated digital footprints in psychiatric research, such as collecting data from different sources, analysis, ethical and research design challenges.

Molecular Psychiatry (2017) **22**, 164–169; doi:10.1038/mp.2016.224; published online 6 December 2016

Psychiatric research conducted in a naturalistic setting involves not only the measurement of mental state – behaviour, feelings and thoughts – but also sampling the contexts of daily life that influence that mental state in social, occupational, and domestic settings.¹ This can provide abundant and invaluable data on environmental contributions to mental illnesses. However, up until now the field of psychiatry has struggled to collect high temporal resolution data from naturalistic settings with high ecological validity using traditional approaches.² This inability to assess the impact of environmental interactions with mental state in real time has hindered progress towards understanding and classifying mental disorders, as well as treating them. With regard to the former, there has been ongoing debate over the validity and adequacy of psychiatric disorder classification systems, for example the Diagnostic and Statistical Manual of Mental Disorders (DSM) and International Statistical Classification of Diseases and Related Health Problems (ICD), that are mostly based on self-reports of symptoms for epidemiology, research, diagnosis and treatment.³ In an attempt to overcome these shortfalls, the US National Institute for Mental Health initiated the Research Domain Criteria project in 2009 to develop a classification scheme combining findings from neuroscience, genetics and cognitive science.⁴ However, biomarkers specific to DSM and ICD categories are yet to eventuate, although there may be subgroups within or across these categories characterised by biomarkers. Psychiatric research in naturalistic settings offers a way to characterise the longitudinal variability of mental disorders, which is a critical shortcoming of cross-sectional studies. This approach may also enable more precise sub grouping of the psychiatric disease

spectrum and in turn open up new and exciting avenues for targeted treatments, as well as better understanding the efficacy of existing treatments.

Expert observations and self-reports – state of the art methods for capturing behaviour and clinical problems in psychiatry – are not suited for collecting high quality temporal data from naturalistic settings representing the subject's real environment. Expert observational methods are associated with procedural problems due to behavioural coding systems, observer biases and reactivity effects as a result of the observation. It is also logistically challenging to observe behaviour in a naturalistic setting over time. As a result, self-reports currently are the most widely used assessment tool within clinical psychiatric research. However, the value of self-reports is reduced by biases,⁵ for example, as a result of individuals being unaware of the cognitions or behaviours in question,⁵ factors such as social desirability of response, the sensitive nature of the construct in question, personality factors, memory processes, or situational characteristics when completing the measure.⁶ Thus, the ability to collect high quality real life data (that are not entirely dependent on self-reports or expert observations in controlled settings) is a critical step to identify sub-groups that might have specific biomarkers and genetic traits.⁷ Can modern technologies, which now permeate everyday life, make it finally possible to characterise psychiatric problems in a naturalistic setting? Digital footprints passively generating high-resolution data from modern technologies have the potential to characterise an individual's behaviour and their environment as well as reflect longitudinal changes.

¹School of Medicine, Flinders University, Adelaide, SA, Australia and ²South Australian Health and Medical Research Institute, Adelaide, SA, Australia. Correspondence: Professor N Bidargaddi, Personal Health Informatics, School of Medicine, Flinders University, Adelaide, SA 5000, Australia.

E-mail: niranjan.bidargaddi@flinders.edu.au

Received 15 May 2016; revised 19 October 2016; accepted 24 October 2016; published online 6 December 2016

WHAT ARE DIGITAL FOOTPRINTS?

'Digital footprints' refer to the data traces arising as a by-product of individuals' day-to-day interactions with mobile and/or Internet-connected technologies⁸ (Figure 1). This paper excludes data that is explicitly collected for health purposes such as electronic health records. With the increasingly ubiquitous ways of accessing the Internet, 3.4 billion individuals (46.4% of the world population) are interacting with the Internet,⁹ primarily through web browsers and software on computers. The interactions encompass a range of daily activities from accessing information, communication, maintaining social connections, online banking and shopping, through services such as Facebook, Twitter, Instagram, LinkedIn, Google and Amazon. As a result of these interactions, an unprecedented passive data set is generated, comprising Internet sites visited, purchase transactions, communication patterns, logs of web searches, purchase records and social media 'likes'. Although the term digital footprint was previously introduced to primarily describe this type of data arising from Internet activity,^{10,11,12} in this paper we broaden the definition to include data arising from day-to-day interactions with newer technologies such as smartphones and other devices.

Interactions with smartphones are further enriching the digital footprints above and beyond data generated from traditional Internet activity. The transition to web enabled smartphones is occurring at a rapid rate, with a reach of 1.9 billion individuals worldwide, including developing regions.¹³ Usually, Individuals mainly interact with their smartphones to communicate, which encompasses phone calls, text messaging or so called Instant Messaging. However, interaction with smartphones also involves the use of web browsers as well as specially designed mobile applications, so called 'apps'. Apps may connect to the Internet directly, but they are also capable of deriving contextual information from mobile phone sensors. Most smartphones have a host of embedded sensors that enable diverse functions.

The advances in miniaturisation of sensors and ubiquitous computing have led to a new type of emerging digital footprint derived from consumer technologies. These include a variety of every day use devices, such as personal monitoring wearables marketed for fitness and wellness purposes, part of the so called Internet of Things. These devices generate high-resolution and

longitudinal behavioural, environmental and physiological data. For example, mattress based sensors can detect chest movement patterns and derive sleep and respiratory rate metrics. Wrist worn devices include accelerometers, wifi, global positioning system and optical sensors to derive activities, locations and physiological measures like heart rate. Following on from these technologies is the development of data collection platforms such as Apple Health Kit and Google Fit, which integrate health data from wearables and fitness apps. Internet of Thing technologies are also resulting in repositories of data on an individual's range of home appliances collected over the Internet for the purpose of automating personalised experiences in the home. For example, Nest, a smart home platform affiliated with Google, collects and interprets data from appliances such as thermostats or smart lights to learn a user's behaviour in order to optimise their experience. What has been demonstrated so far utilising digital footprints?

DIGITAL FOOTPRINTS USED IN THE COMMERCIAL SECTOR AND HEALTH RELEVANT TO PSYCHIATRIC RESEARCH

Extensive research on harnessing digital footprints in the commercial sector exists in Marketing and Consumer Behaviour field. It ranges from nuanced understanding of factors that influence customers spending behaviour, determining effectiveness of advertisements strategies, and tailoring strategies that enhance the experiences of customers. For example, analysis of online shopping data suggested that repeat customers on online shopping website will buy more products per order over time.¹⁴ Thus, even for low margin purchases it is valuable for a company to retain customers over a long period. In shop location data has also been used to study spending behaviour within high-street shops. A study that offered mobile promotions for products via phone and collected customer's location data when they in store using radio-frequency identification technology showed that unplanned spending increases when targeted mobile promotion aims at increasing store-path length.¹⁵ Such data has also been mined at population level to predict emerging market opportunities as noted in analysis of real time location data of several million users to accurately predict number of sales when Apple

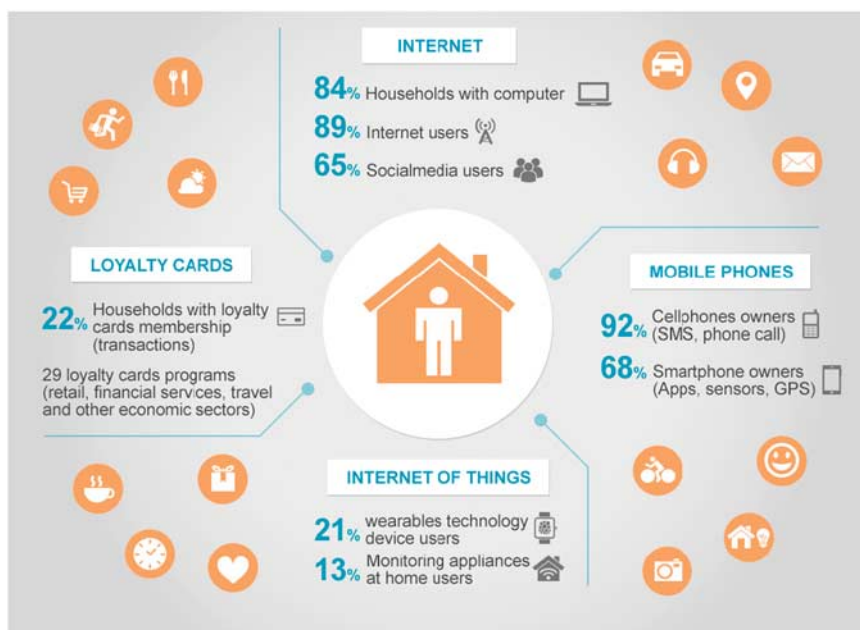


Figure 1. Every day interaction with technologies: prevalence of key technologies in United States.⁴⁵⁻⁵⁰

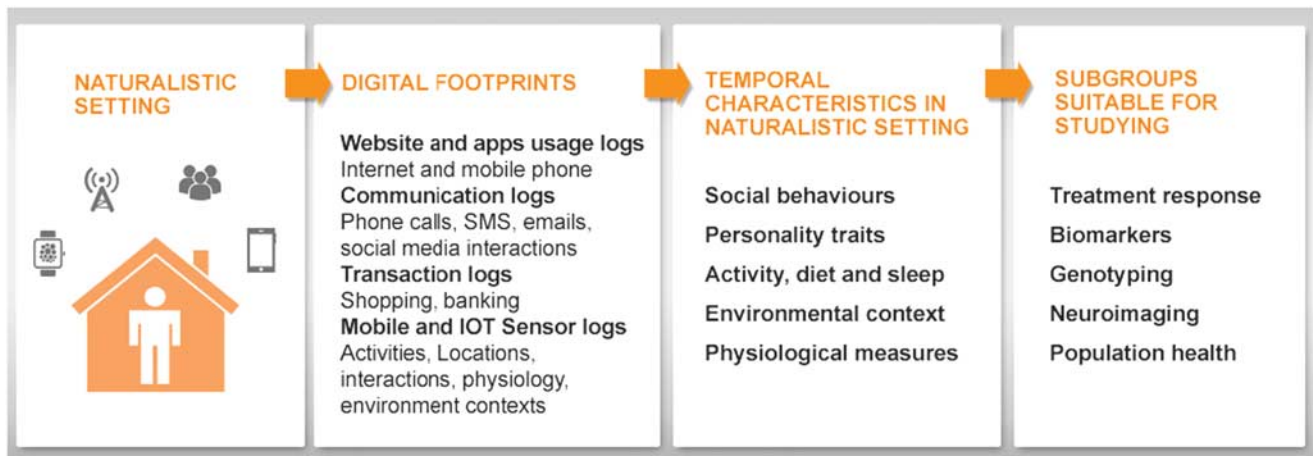


Figure 2. Digital footprints from every day technologies and their uses.

release iPhone or McDonald's introduced Happy Meals even before they occurred.¹⁶ In summary there is compelling evidence in the marketing field that data points arising from digitisation of everyday life can be used to offer personalised products suited to individuals behaviour and preferences.¹⁷ Can this approach be similarly adopted to identify well-characterised subgroups of relevance to psychiatric research? Although research considering and leveraging digital footprints in their entirety is lacking, several studies have demonstrated their utility in deriving various attributes, which are temporal, obtained from naturalistic environment and highly relevant to psychiatric phenotypes (Figure 2).

It has been demonstrated that social behaviour and personality traits, both attributes highly relevant to psychiatric research, can be derived from past email and social media activity. Large scale analysis of social network posts from University students showed that individuals affected by hurricanes are more likely to strengthen interactions while maintaining the same number of friends than unaffected individuals.¹⁸ Email analysis of 1000 individuals' email logs over 4 years in a work environment revealed individuals have unique patterns of ties within their social network which endure over extended period of time.¹⁹ For example, social signatures, that is individuals' efforts for maintaining social relationships, are stable over several years, as are communication strategies. 'Likes' are means by which users can identify their own preferred Internet sites and interests. In a large US study with over 58 000 individuals, Facebook 'likes' predicted the Big Five personality traits as assessed with the NEO-PI-R with reasonable accuracy, which increased with the analysis of more 'likes'.¹⁰

Of greater immediate relevance to psychiatry are studies in which health parameters, symptoms, environmental contexts or risk factors for mental disorders have directly been derived from digital footprints. A recent UK study analysed a large data set of supermarket shopping records from 24 879 households in the UK to reveal insights between shopping behaviour and health.²⁰ For example, visiting store chains more frequently correlated with higher fruit and vegetable consumption. Similarly, temporal characterisation of sleep and activity behaviours along with the location data can be obtained from devices that individuals are already using. A recent analysis of such data from users of the fitness tracker Jawbone UP revealed how different urban environments influence sleep and physical activity behaviour and suggested that students at higher ranking universities have later bedtimes, but do not necessarily sleep less than students at lower ranking universities.²¹

Digital footprints not only facilitate research in a naturalistic setting, but also describe this setting in greater detail. Several studies have demonstrated digital footprints can provide real time data on individual's environment, that when linked with mental health data, may contribute towards our understanding of the aetiology of psychiatric problems. For example, the Mappiness study from the London School of Economics²² sampled mood from over 40 000 volunteers through its mobile app and correlated it with environmental features derived from 20 000 photographs that individuals already had in their phone to demonstrate how people's feelings are affected by characteristics of their current environment. When temporal location data is collected and analysed a population level, the patterns within human behaviour itself can be uncovered.²³ These examples for deriving contexts for technology use are not confined to the developed world. Research analysing mobile phone metadata, such as volume and frequency of use, of a large sample in Rwanda demonstrated that information on poverty or wealth can be derived.²⁴

Already direct applications of these approaches in psychiatry are appearing in the literature, particularly in situations where the patterns of technology use itself has pathological characteristics.²⁵ In particular, there is a well-established association between pathological Internet use ('Internet addiction') and attention deficit/hyperactivity disorder and obsessive-compulsive symptomatology.²⁶ Similarly, another application has emerged in the discovery that deviations from regular patterns of digital footprints can be predictive of mental illness. A recent study in patients with bipolar disorder found that reduced activity, as assessed via cell tower movements, was associated with both increased depressive and manic symptoms.²⁷

POINTS TO BE CONSIDERED

This paper introduces digital footprints, data generated passively from day to day interaction with Internet technologies, as a means to facilitate psychiatric research in a naturalistic setting. Utilizing digital footprints for refining diagnostic criteria or monitor treatment response is a relatively new area. Promisingly, recent studies have demonstrated the suitability of data from mobile phones to derive several behavioural indicators of mental health²⁸ and longitudinal phenotype characterisation in psychiatry,²⁹ and the feasibility of collecting such data from psychiatric patients.^{6,30,31} These early studies have not tapped into the breadth of digital footprints as they only focused on specific variables or devices. Furthermore they have adopted purposeful data collection approaches by providing dedicated devices to

participants instead of harnessing data from technology individuals are already using, which would constitute a more naturalistic setting and is also scalable. Realizing the potential of digital footprints requires addressing challenges discussed below.

Data collection from different sources

Nearly all of the services and application that collect and store such footprints are outside the mainstream of traditional health care or public health research. Often individuals' despite being the source of the data have very little understanding and control of their digital footprints, and also benefit the least. There is a need to develop and maintain knowledgebase that maintains a catalogue of different technology sources, characteristics of digital footprints they create, ways to access them guided by relevant ethical, legal, cultural, geographical and regulatory issues.³² It needs to be agile, responsive to changing technology landscape and be beneficial to both industry and research communities.

Untangling the ownership of digital footprints is a unique challenge related to collection of individual's digital footprints from identified sources. Since digital footprints may be a by-product of an individual's transaction in a commercial environment, data are owned by businesses. One model is to give users control over their own data and equip them with capabilities to donate and share it as required for their benefits. Public discussions and ideas on shifting control of such data back to individuals is an emerging theme. For example, consumer empowerment initiatives such as the 'midata' program adopted by several businesses in the UK are now allowing individuals access to their personal consumption and transaction data so that they can use it to find alternative services.³³ In some instances, change in mode of business is also facilitating better data access. For example, users are gaining access and control to their own transaction logs by virtue of online banking and shopping. Another impetus for data access is maintaining competitiveness in commercial space by creating an ecosystem of users between synergistic products, which requires sharing of data between services. Given the rapidly evolving technology landscape, unless there is a financial and business incentive, businesses are unlikely to invest into effective data provision processes. A financial market for personal data where different solutions such as Citizens data cooperative³⁴ and Health data banks³⁵ could accelerate access to personal data.

Parallely, industry wide adoption of infrastructure that enables access to digital footprints standardised and consistent manner is equally critical. Increasingly most commercial services are offering data through application programming interfaces (APIs), special publically available software that can be integrated with other applications to directly exchange needed data based on selected criteria. For example, over 5500 banks use Open Financial Exchange, which is a standard to exchange financial data between them. Adoption of such standards has allowed individuals access to their transactions data over the Internet. Large Internet services like Amazon, Facebook, Fitbit, Google and Apple make data already available through APIs in some form. However, data is often released in a restricted manner, rarely complete or raw data is provided and different taxonomies are used. Hence, there is a need not only to design technological solutions that facilitate access to personal data but also to promote a practice that facilitates access to high integrity data.

Another approach for collecting digital footprints at population level is by adopting successful strategies and processes used by Internet data aggregators.⁸ Commercial data aggregators such as BlueKai, OutBrain and Rio have partnered with major web content publishers to deploy tracking technologies which has resulted in a longitudinal cohort of over 100 million users with several data points revealing demographics and online behaviours preferences over time.⁸ This model of collaboration between web site

publishers and data aggregators in collection and harvesting of data using third-party cookies could equally be relevant to collecting and curating data for research purposes.

Addressing analysis challenges

An individual's digital footprints is made up of several data points over time, with differing data types and temporal resolutions, by the nature and usage patterns of Internet connected applications and devices. The nature of the data poses several challenges with conventional analysis approaches, which work well with perfectly sampled data collected from controlled experiments. The first major hurdle is related to suitability of digital footprints for comparison between individuals, when there is a variance in technology use. This requires correcting inter-individual variances arising from different appliances producing same measurements in the analysis. Validation studies of different sources and appliances can help address this issue by providing norms that account for inter-individual differences, as well as variances between appliances producing same data.³⁶ Creating open source repositories of data from well-characterised populations can further assist with cross comparison with newer technologies and establishing norms. Second, at an individual level the varying temporal resolution observed for each measurement as is common with data from smartphone sensors or wearables may not just be missing data, but be indicative of behaviour change or the lack of thereof. One approach for analysing this type of data involves developing reference data sets from well-characterised patient groups and applying machine learning and signal processing techniques, such as Markov models. Using a reference data set, these techniques can be applied to first derive a range of features from patchy data points of varying resolution and subsequently train models that describe states of behaviours and mood.³⁷ The trained models can be deployed to identify similarities and differences between individuals as well as detects changes over time from their digital footprints. Turning data into features and then states of behaviour addresses not only the challenges of varying resolution but the characteristics of varying resolution such as missing data or patterns of usage can itself be treated as feature. A study using Markov models on mood ratings obtained from students identified two latent classes of mood regulation patterns, which could be predicted with high accuracy.³⁸ It has been recognised recently that these analysis approaches are increasingly becoming common in psychiatry³⁹ and will facilitate the inclusion and analysis of digital footprints in future research.

Rethinking study design

To collect and analyse digital footprint for health research, research design challenges associated with this field need to be taken into consideration. Data collection takes place in a natural environment and at, or close to, the time of events occurring and thus outside the control of researchers. Digital footprints can be considered as a by-proxy observational method to assess naturalistic behaviour, as well as environments, while overcoming many problems traditionally associated with behavioural observations.⁴⁰ The ubiquity of digital technology in our lives removes many logistical barriers, and problems such as loss of information due to the limited capacity of the observer or coding systems can be mitigated. As a result, behaviours with very low base rates can be easily studied. Particularly within Research Domain Criteria projects⁴ digital footprints can be incorporated into study methodology to obtain ecologically valid data for difficult-to-assess constructs. Data collection in digital footprints itself is less prone to biases, as the process of collecting data from everyday interactions with technology is unlikely to induce reactivity (although this may be desired, as in the context of, for example, wearable activity monitors). Digital footprints can be

considered as an extended phenotype, a term described by Richard Dawkins,⁴¹ and therefore in part an external manifestation of biological processes or genotypes. However, depending on the constructs studied and technologies analysed it is important to acknowledge the potential for sampling biases in digital footprints. There are differences in the penetration of technologies generating digital footprints with regard to age,⁴² ethnicity,⁴³ country, socio-economic status and other variables. Although, it is likely for this 'digital divide' to close over time as technologies become more affordable and widespread, it is important to consider when analysing and interpreting digital footprints.

Ethical challenges

Utilising data from commercial sources that were not purposely collected with prior consent for mental health raises ethical concerns. Such information has been widely used in the commercial sector for marketing purposes, very often with low standards of privacy and ethics. While consumer attitudes are often negative with regard to the use of their data for commercial purposes, they are surprisingly tolerant for use in research or for the benefit of other sufferers.⁴⁴ Learning from and adapting strategies used by marketers while adhering to ethical standards of research could lead to enormous progress in the field of psychiatric research, as well as for health research in general. Developing strategies that offer individuals better control of their diverse digital footprints with opportunities to control the information they wish to share can balance the risk and benefits of such data.

CONCLUSIONS

The Internet, emails, text messaging, social media, smartphone apps, wearable technology and the Internet of Things are now successively permeating every aspect of peoples' lives. And the by-product of our daily digital activities – digital footprints – offers an exciting opportunity for groundbreaking psychiatric research, including a novel approach for the collection of real-time data on environmental contributors to mental illness. Psychiatric disorders are the outcome of gene-environment interactions. In parallel to the recent progress in genomics, digital footprints can now substantially expand the assessment of environmental data. Altogether, this will substantially increase our ability to dissect gene-environment interactions in mental illnesses at an unprecedented level. Marketing technologies have finessed the accumulation of knowledge from digital footprints, and demonstrated an exploitable connection to the future behaviour of an individual consumer. We propose to harness the same connection for psychiatric research: by observing digital footprints on an epidemiological scale and in naturalistic settings, new scientific knowledge can be created that will enable us to develop and apply more effective treatments and public health policies in the future.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- 1 Bellack AS, Hersen M. Handbook of Behavior Therapy in the Psychiatric Setting [Internet]. Springer US: Boston, MA, 1993. Available at: <http://link.springer.com/10.1007/978-1-4899-2430-8>.
- 2 Markowitz A, Błaskiewicz K, Montag C, Switala C, Schlaepfer TE. Psycho-Informatics: big data shaping modern psychometrics. *Med Hypotheses* 2014; **82**: 405–411.
- 3 Phillips J, Frances A, Cerullo MA, Chardavoyne J, Decker HS, First MB et al. The six most essential questions in psychiatric diagnosis: a pluralogue part 1:

conceptual and definitional issues in psychiatric diagnosis. *Philos Ethics Humanit Med* 2012; **7**: 3.

- 4 Cuthbert BN. The RDoC framework: facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. *World Psychiatry* 2014; **13**: 28–35.
- 5 Nisbett RE, Wilson TD. Telling more than we can know: verbal reports on mental processes. 1977 [cited 8 February 2016]; Available from <http://deepblue.lib.umich.edu/handle/2027.42/92167>.
- 6 Ben-Zeev D, Scherer EA, Wang R, Xie H, Campbell AT. Next-generation psychiatric assessment: Using smartphone sensors to monitor behavior and mental health. *Psychiatr Rehabil J*. 2015; **38**: 218–226.
- 7 McInnis MG, Greden JF. Longitudinal studies: an essential component for complex psychiatric disorders. *Neurosci Res* [Internet]. [cited 18 January 2016]; Available from <http://www.sciencedirect.com/science/article/pii/S0168010215001522>.
- 8 Digital Footprints - An Internet Society Reference Framework [Internet]. [cited 4 February 2016]. Available from <http://www.Internetsociety.org/sites/default/files/Digital%20Footprints%20-%20An%20Internet%20Society%20Reference%20Framework.pdf>.
- 9 World Internet Users Statistics and 2015 World Population Stats [Internet]. [cited 18 April 2016]. Available from <http://www.Internetworldstats.com/stats.htm>.
- 10 Kosinski M, Stillwell D, Graepel T. Private traits and attributes are predictable from digital records of human behavior. *Proc Natl Acad Sci* 2013; **110**: 5802–5805.
- 11 Youyou W, Kosinski M, Stillwell D. Computer-based personality judgments are more accurate than those made by humans. *Proc Natl Acad Sci* 2015; **112**: 1036–1040.
- 12 Gittelman S, Lange V, Gotway Crawford CA, Okoro CA, Lieb E, Dhingra SS et al. A New Source of Data for Public Health Surveillance: Facebook Likes. *J Med Internet Res* [Internet]. [cited 11 January 2016];17(4), 2015. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4419195/>.
- 13 GSMA. The Mobile Economy 2015 [Internet]. [cited 27 April 2016]. Available from http://www.gsma.com/mobileeconomy.com/GSMA_Global_Mobile_Economy_Report_2015.pdf.
- 14 Reibstein DJ. What attracts customers to online stores, and what keeps them coming back? *J Acad Mark Sci* 2002; **30**: 465–473.
- 15 Hui SK, Inman JJ, Huang Y, Suher J. The effect of in-store travel distance on unplanned spending: applications to mobile promotion strategies. *J Mark* 2013; **77**: 1–16.
- 16 Glueck J. McDonald's vs Taco Bell: Who's winning the breakfast wars? – Four-square Direct [Internet]. Medium. 2015 [cited 3 September 2016]. Available at: <https://medium.com/foursquare-direct/mcdonald-s-vs-taco-bell-who-s-winning-the-breakfast-wars-a27f11e58483#.7twa5Zopoc>.
- 17 Salonen V, Karjaluoto H. Web personalization: the state of the art and future avenues for research and practice. *Telemat Inform* 2016; **33**: 1088–1104.
- 18 Phan TQ, Airoldi EM. A natural experiment of social network formation and dynamics. *Proc Natl Acad Sci* 2015; **112**: 6595–6600.
- 19 Godoy-Lorite A, Guimera R, Sales-Pardo M. Long-term evolution of email networks: statistical regularities, predictability and stability of social behaviors. *PLoS ONE* 2016; **11**: e0146113.
- 20 Pechey R, Monsivais P. Supermarket Choice, Shopping Behavior, Socioeconomic Status, and Food Purchases. *Am J Prev Med*. 2015; **49**: 868–877.
- 21 The Jawbone Blog - [Internet]. The Jawbone Blog. [cited 5 May 2016]. Available from <https://jawbone.com/blog>.
- 22 Bryson A, MacKerron G. Are you happy while you work? *Econ J* 2016; doi:10.1111/eoj.12269.
- 23 Pappalardo L, Simini F, Rinzivillo S, Pedreschi D, Giannotti F, Barabási A-L. Returners and explorers dichotomy in human mobility. *Nat Commun* 2015; **6**: 8166.
- 24 Blumenstock J, Cadamuro G, On R. Predicting poverty and wealth from mobile phone metadata. *Science* 2015; **350**: 1073–1076.
- 25 Nathan D, Shukla L, Kandasamy A, Benegal V. Facebook role play addiction – a comorbidity with multiple compulsive-impulsive spectrum disorders. *J Behav Addict* 2016; **5**: 373–377.
- 26 Carli V, Durkee T, Wasserman D, Hadlaczky G, Despalins R, Kramarz E et al. The association between pathological Internet use and comorbid psychopathology: a systematic review. *Psychopathology* 2013; **46**: 1–13.
- 27 Beiwinkel T, Kindermann S, Maier A, Kerl C, Moock J, Barbian G et al. Using smartphones to monitor bipolar disorder symptoms: a pilot study. *JMIR Ment Health* 2016; **3**: e2.
- 28 Glenn T, Monteith S. New measures of mental state and behavior based on data collected from sensors, smartphones, and the Internet. *Current Psychiatry Reports* [Internet]. cited Oct 29 2015;16(12), 2014. Available at: <http://link.springer.com/10.1007/s11920-014-0523-3>.
- 29 Wenzel M, Kubiak T, Ebner-Priemer UW. Ambulatory assessment as a means of longitudinal phenotypes characterization in psychiatric disorders. *Neuroscience Research*[Internet]. [cited 15 January 2016]; Available at: <http://www.sciencedirect.com/science/article/pii/S0168010214002417>.

- 30 Saeb S, Zhang M, Karr CJ, Schueller SM, Corden ME, Kording KP *et al*. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study. *J Med Internet Res* 2015; **17**: e175.
- 31 Schwartz S, Schultz S, Reider A, Saunders EFH. Daily mood monitoring of symptoms using smartphones in bipolar disorder: a pilot study assessing the feasibility of ecological momentary assessment. *J Affect Disord* 2016; **191**: 88–93.
- 32 Trust W. Digital Phenotypes – Health research in the digital age [Internet]. Wellcome Trust Blog. 2016 [cited 5 August 2016]. Available at: <https://blog.wellcome.ac.uk/2016/03/08/digital-phenotypes-health-research-in-the-digital-age/>.
- 33 Midata for personal current accounts [Internet]. [cited 14 January 2016]. Available from <https://www.pcamidata.co.uk/>.
- 34 Hafen E, Kossmann D, Brand A. Health data cooperatives - citizen empowerment. *Methods Inf Med*. 2014; **53**: 82–86.
- 35 Shabo A. It's time for health record banking! *Methods Inf Med* 2014; **53**: 63–65.
- 36 Case MA, Burwick HA, Volpp KG, Patel MS. Accuracy of smartphone applications and wearable devices for tracking physical activity data. *J Am Med Assoc* 2015; **313**: 625–626.
- 37 Eagle N, (Sandy) Pentland A. Reality mining: sensing complex social systems. *Pers Ubiquitous Comput*. 2006; **10**: 255–268.
- 38 Crayen C, Eid M, Lischetzke T, Courvoisier DS, Vermunt JK. Exploring dynamics in mood regulation—mixture latent Markov modeling of ambulatory assessment data. *Psychosom Med* 2012; **74**: 366–376.
- 39 Passos IC, Mwangi B, Kapczynski F. Big data analytics and machine learning: 2015 and beyond. *Lancet Psychiatry* 2016; **3**: 13–15.
- 40 Lipinski D, Nelson R. Problems in the use of naturalistic observation as a means of behavioral assessment. *Behav Ther* 1974; **5**: 341–351.
- 41 Dawkins R. The extended phenotype: the long reach of the gene [Internet]. Oxford Paperbacks; 1999 [cited 27 April 2016]. Available from <https://books.google.com/books?hl=en&lr=&id=vsIZAgAAQBAJ&oi=fnd&pg=PP2&dq=%22said+to+act.+I+happen+to+be+an+ethologist,+but+I+hope+preoccupations%22+%22to+enjoy+a+professional+book+as+a+spectator,+even+if+not+a%22+&ots=lzxtVTABDn&sig=9YAyUoNlnRz6dmdL8SZdcsEvuo>.
- 42 Zickuhr K, Madden M. Older adults and Internet use [Internet]. Pew Research Center: Internet, Science & Tech. 2012 [cited 3 September 2016]. Available from <http://www.pewInternet.org/2012/06/06/older-adults-and-Internet-use/>.
- 43 Ennis L, Rose D, Denis M, Pandit N, Wykes T. Can't surf, won't surf: the digital divide in mental health. *J Ment Health Abingdon Engl* 2012; **21**: 395–403.
- 44 Jackson SL, Mejilla R, Darer JD, Oster NV, Ralston JD, Leveille SG *et al*. Patients who share transparent visit notes with others: characteristics, risks, and benefits. *J Med Internet Res* 2014; **16**: e247.
- 45 U.S. Customer Loyalty Program Memberships Top 3 Billion For First Time, 2015 COLLOQUY Census Shows [Internet]. COLLOQUY. [cited 8 May 2016]. Available from: [/latest-news/](http://latest-news/).
- 46 Monica Anderson. Technology Device Ownership: 2015 [Internet]. Pew Research Center: Internet, Science & Tech. 2015 [cited 8 May 2016]. Available from <http://www.pewInternet.org/2015/10/29/technology-device-ownership-2015/>.
- 47 Rainie L comments DC. Census: Computer ownership, Internet connection varies widely across U.S. [Internet]. Pew Research Center. 2014 [cited 8 May 2016]. Available at <http://www.pewresearch.org/fact-tank/2014/09/19/census-computer-ownership-Internet-connection-varies-widely-across-u-s/>.
- 48 Andrew Perrin. Social Media Usage: 2005–2015 [Internet]. Pew Research Center: Internet, Science & Tech. 2015 [cited 8 May 2016]. Available at: <http://www.pewInternet.org/2015/10/08/social-networking-usage-2005-2015/>.
- 49 Internet Users by Country (2016) - Internet Live Stats [Internet]. [cited 8 May 2016]. Available at <http://www.internetlivestats.com/Internet-users-by-country/>.
- 50 Smart Home Ecosystem: IoT and Consumers [Internet]. [cited 8 May 2016]. Available at: <http://www.parksassociates.com/whitepapers/iot-smart-devices>.