Medical Hypotheses 82 (2014) 405–411

Contents lists available at ScienceDirect

Medical Hypotheses

journal homepage: www.elsevier.com/locate/mehy

Psycho-Informatics: Big Data shaping modern psychometrics

Alexander Markowetz [a], Konrad Błaszkiewicz [a], Christian Montag [b], Christina Switala [c], Thomas E. Schlaepfer [c,*]

[a] Department of Computer Science, University of Bonn, Germany
[b] Department of Psychology, University of Bonn, Germany
[c] Department of Psychiatry and Psychotherapy, University of Bonn, Germany

ARTICLE INFO

Article history:
Received 6 September 2013
Accepted 19 November 2013

ABSTRACT

For the first time in history, it is possible to study human behavior on great scale *and* in fine detail simultaneously. Online services and ubiquitous computational devices, such as smartphones and modern cars, record our everyday activity. The resulting Big Data offers unprecedented opportunities for tracking and analyzing behavior. This paper hypothesizes the applicability and impact of Big Data technologies in the context of psychometrics both for research and clinical applications. It first outlines the state of the art, including the severe shortcomings with respect to quality and quantity of the resulting data. It then presents a technological vision, comprised of (i) numerous data sources such as mobile devices and sensors, (ii) a central data store, and (iii) an analytical platform, employing techniques from data mining and machine learning. To further illustrate the dramatic benefits of the proposed methodologies, the paper then outlines two current projects, logging and analyzing smartphone usage. One such study attempts to thereby quantify severity of major depression dynamically; the other investigates (mobile) Internet Addiction. Finally, the paper addresses some of the ethical issues inherent to Big Data technologies. In summary, the proposed approach is about to induce the single biggest methodological shift since the beginning of psychology or psychiatry. The resulting range of applications will dramatically shape the daily routines of researches and medical practitioners alike. Indeed, transferring techniques from computer science to psychiatry and psychology is about to establish Psycho-Informatics, an entire research direction of its own.

Introduction

Reliable measurements of emotion, cognition and behavior are of equal and central importance to psychiatry and psychology. Despite the crucial role of these parameters, the basic methodology has remained essentially unchanged for the better half of a century. To this day, researchers rely on clinicians' observations and self-rated psychometric tests. Requiring specially trained experts for conducting interviews or observations, these methods found widespread use in research, but failed to penetrate clinical practice. And, despite the highly skilled evaluators (and associated cost), the gathered data remains rather unsatisfactory, both in terms of quantity *and* quality. As outlined below, these shortcomings are inherent to the methodologies themselves, and cannot be overcome by enlarging financial budgets. At the same time, there appears to be no immediate successor. On the one side, novel and scientifically interesting methods like neuroimaging and genetics represent promising approaches to evaluate the course of psychiatric treatment [1]. On the other side, these approaches are still in an early stage of development, costly and often do not provide reliable diagnostics for individual patients.

The temporal granularity at which traditional methods collect data commonly is too coarse to reveal fine-granular patterns. The most important drawbacks are of entirely practical nature: around-the-clock shadowing is neither affordable nor acceptable to a participant. Instead, researchers employ questionnaires at fixed intervals, essentially relying on the participants' self-report. This method naturally imposes a bound on the temporal granularity at which participants can be interviewed: weekly, monthly, or at an even coarser level. Furthermore, it is not possible to execute the identical psychometric test multiple times over the course of a single day. Memory and training effects would limit the reliability of the ratings. Next, holding interviews at high frequency would be prohibitively expensive, because such interviews have to be conducted by a trained professional. Depending on the level of training

* Corresponding author. Address: Departments of Psychiatry and Psychotherapy/University Hospital, Sigmund-Freud-Strasse 25, 53105 Bonn, Germany. Tel.: +49 228 287 14715; fax: +49 228 287 15025.

E-mail address: schlaepf@jhmi.edu (T.E. Schlaepfer).

http://dx.doi.org/10.1016/j.mehy.2013.11.030
0306-9877/

and employment status, a single interview quickly costs several hundred Euros. In addition, the necessary appointments impose too great of a burden on the participant, especially when the content of the interviews only relies to negative aspects of life such as psychopathological disorders. Including travel, a single interview can consume the better half of a day, a burden that only be imposed infrequently (in particular with participants pursuing a professional career). Self-reports in form of diaries do not provide a viable solution either. This method, too, quickly meets a limit of how much time commitment can be expected from a participant. In sum, the constraints (i) reduce the temporal granularity at which data can be gathered, and (ii) pose tremendous problems for longitudinal studies with respect to the amount and completeness of data gathered over an extended time range.

Unfortunately, data collected by the traditional means is also strongly biased. Most notably, it is commonly faulty and distorted, due to poor recollection of the variable of interest. This holds especially for coarse intervals of reporting, and especially for questions regarding interaction with digital devices. Very few people could accurately report how often they have checked their email over the past 10 days (which would be an interest variable to study Internet use/addiction). Additionally, reports about variables from other research areas, such as subjective well-being, tend to simply reflect altered psychological states. In particular, it has been shown that people use their momentary affective state for judging how happy and satisfied they are with their lives in general. A depressed patient for example will usually see his/her well-being, social functioning, and living conditions worse than they would appear to an independent observer, or to himself/herself after recovery [2]. Thus, self-reports are affected by the state of mind at time of reporting, and the social desirability of the reported behavior. Together, these factors introduce significant noise to infrequently recorded data. Clinician-rated psychometric tests, entail the risk of being similarly biased, since assessments of experts are not entirely objective. In this context, the term "objective" assessment is misleading and should be replaced by "external", as this evaluation might reflect the subjective view of the assessor himself [2].

In short, data gathered by traditional means thus capture the situation of a study's participant or patient rather poorly. It is too coarse to show temporal patterns, and generally lacks dynamics. Additionally, it commonly employs shallow scales, thus quickly encounter floor effects. Most questionnaires regarding depression, for example, only permit answers on each item on a scale of 0–3 [3–5]. The effects of these coarse measurements are dramatic, because novel psychotropic substances frequently become stuck during the development phase, because (visible) positive effects cannot be quantified reliably [6,7]. Clearly, innovation in methodology has long been overdue.

As early as the seventies, researchers circulated the idea of *actigraphy* as a simple and non-invasive method for monitoring human rest and activity cycles. Inter alia, they measured sleep patterns [8,9] and circadian rhythms [10] via specific *actimetry sensors*, worn on the body of the patient. While this approach overcame some of the obstacles faced by questionnaires, it did not quite hit the mark. Early technology was rather simple, rendering sensors complex, expensive and socially awkward, thus requiring substantial compliance and discipline from the patient. In some areas of research such as neuropharmacology, actigraphy only was administered in very few cases [11]. In recent years, miniaturization of digital devices has given new rise to the methodology. Sensors have become smaller, less power-hungry, and can independently transmit their data. While finally practical, the central obstacle to actigraphy remains: the patient/participant must be coaxed into carrying a sensor for a substantial period of time. In sum, actigraphy has only been used sporadically in most areas of psychiatry and psychology. Due to miniaturization, it is about to enjoy a second lease of life, but will ultimately be made redundant by sensor-less methods of tracking.

In this paper, we propose observing behavior directly on digital devices and services, such as laptops, social networks, or even cars. Specifically, we focus on user interaction with smartphones. Carried on the person, around the clock, and used for a wide range of (informal) communication, these devices constitute a particularly rich and intimate source of information. The gathered data is of highest quality, gathered entirely in the background, and automatically forwarded to a central server. The method thus burdens neither patient/participant nor researcher. Most importantly, avoids the dominant sources of bias, commonly encountered by self-reports and questionnaires.

For several areas of research, the proposed methodologies constitute the only viable solution. Most notably, it constitutes the only valid measure for usage and abuse of digital media. Kimberly Young [12] first saw a problem for the human condition when excessively using the Internet, an issue also put forward for the usage of mobile phones [13]. Whether the observed phenomena constitute a 'new disorder' is a matter of heated debate [14]. Although excessive use of the Internet is not a distinct disorder in the DSM-V, evidence from both psychology, psychiatry and the neurosciences suggest that "Internet addiction" constitutes a substantial challenge [15,16]. While a high daily "dosage" does not qualify for an addiction, a rising number of hours spent with the phone over a certain time could indicate developing tolerance. In any case, such behavior must be recorded directly on the device. Ordinary patients/participants cannot be expected to accurately answer how often they unlock their phone each day (up to 200 times, according to our preliminary experimental findings). The particularly poor recollection in this context arises due to the "virtual" character of phone behavior. Alcohol consumption for example, is significantly easier to quantify, if only by the number of empty bottles.

The proposed methodology is about to equally revolutionize the work of researchers with more classic research agendas, such as personality or behavior. Recently, Kosinski et al. impressively inferred personality traits from the behavior on the Internet platform Facebook [17]. Yet, such research endeavors only mark the beginning of tight collaboration between psychology/psychiatry and informatics. After all, Facebook usage 'only' represents a rather narrow glimpse on people's lives. By comparison, how much can we learn about the human condition when monitoring mobile phones 24 h/7 days a week? The socially outgoing (extraverted) person could easily be detected by the amount of in- and outcoming calls, indicating a large active social network. The introverted person in contrast might display longer reading sessions, perhaps using an e-book application. The person being open for new experiences (another of the Big Five Factors of Personality describing human characteristics by McCrae and John [18]) might often install and test new apps. Numerous such dependent variables can be detected by observing humans through their mobile phone interactions. These measures will capture the human condition more precise than ever. For the first time, psychiatrists and psychologists can observe human behavior on a large scale, in the finest temporal granularity. They can thus assess the course of treatment and disease in a temporal continuum, instead of relying on selective snapshots.

Equally, the proposed methodology is about to revolutionize clinical therapy, a role in which it will affect our everyday lives to an even higher degree. In this scenario, patients track a wide range of personal data, from phones, cars, and fridges. From this raw (and rather cryptic) data, large-scale analysis extracts meaningful indices, such as an "activity index", or a "social interaction index". The patient can then self-track his condition. He is reassured that it is not worsening. Or, if a worsening of his health condition occurs, he could confidently ask for an ad-hoc appointment with his doctor. In addition, he can explore interdependencies between his health condition and his lifestyle, such as staying up late, or working out. Most importantly, he can provide (selected) data access to his coach, therapist or doctor.

For the clinician, this methodology enables an entire range of new options. For the first time, he does not have to rely on the (poor) self-report of his patient. Instead, he receives clear indicators of the patient's mental state, and changes therein, in a fine-granular temporal resolution. He can thus observe the continuous changes of health parameters over time (to follow the course of a disease, or the progress of therapy). The clinician will also be able to investigate changes of his patient throughout the day, and fine-tune timing and dosage of medication, providing a highly individualized therapy. For example, he thus could match medication doses in a patient suffering from schizophrenia. The clinician can even prescribe a range of dosage, from which the patient can independently choose, according to his or her latest data. The therapist can be automatically alarmed when symptom data indicates a critical situation. In this case, he can intervene via phone, video conference, or an ad-hoc appointment. At the same time, regular appointments can be spaced further apart.

Most importantly, the proposed methodology is significantly cheaper than personal interaction with a therapist. This profane observation has vast implications, opening the application area towards wellness and prevention for large amounts of people. Currently, society focuses its limited therapeutic resources on sick patients. In the future, data driven early warning systems will enable us to help people a long time before their conditions becomes serious or chronic. Raising red flags early, some people might just need to attend a seminar on sustainable usage of digital media, or an extended vacation, or the HR department talk to their chaotic manager. Eventually, most corporations will deploy data driven preventive mental health programs. The ethical perspective (as discussed below) only constitutes a fraction of the challenges these services face. The integration into the processes and structures of large corporations might turn out far more difficult. Yet, occupational doctors can serve as a blueprint for a data driven occupational mental health service, leading to their widespread deployment much sooner than anticipated.

The remainder of this paper is structured as follows. Next, we outline the underlying technological vision, comprised of various data sources, and means to store and analyze the data. Subsequently, we introduce two current studies and respective hypotheses. One study tracks depression, the other investigates the misuse of mobile phones. We then touch upon the ethical aspects of the proposed methodology, a topic we feel very strongly about. The article ends with an outlook on the anticipated changes in research and therapy. As we outline, the proposed methodology will shape, if not revolutionize, psychiatry and psychology. The envisioned shift will be massive, touch every aspect of both sciences, and eventually create its own field of research: Psycho-Informatics.

Underlying technological vision

This paper's is based on a single central thesis. The user's mental state, we claim, affects the way he interacts with a machine. A stressed user may thus generate more typographic errors than ordinarily; a depressed user may communicate less over his phone than previously. Conversely, so the claim continues, changes in his interaction with a machine reflect changes in his mental state. Modern computer science enables us to automatically gather the appropriate data, transfer, and analyze it, all at very little cost. The proposed methodology, so we hypothesize, might outperform traditional methods in both data quality as well as quantity.

The twenty-first century will likely be remembered as the age of *Big Data*. Recent advancements in hard- and software enable us to store and analyze massive amounts of data, at surprisingly little cost. Currently, such technology is most prominently employed by search engines, social networks, credit card issuers and insurance companies. These very different businesses gather massive amounts of data, e.g., to seek underlying patterns or assign scores to individual users. Based on these quantitative measures, they automatize decision-making processes, such as which advertisement to show, whether to approve a transaction, or to grant a credit line. In the context of psychology and psychiatry, we propose Big Data to produce psychometric parameters and to trace the course of a disorder.

Unfortunately, terms like "Big Data" or "data mining" are surrounded by a significant amount of buzz. Frequently fed by marketing departments from IT companies, the hype often obfuscates the actual potential and limits. Yet, as this section should illustrate, the potential of these technologies is indeed immense. To complicate the terminology further, there is a significant overlap between the areas and concepts. Terms such as "data mining" and "Big Data" are often used interchangeably, even by experts. We thus next clarify the core concepts, before outlining their applicability in the context of psychiatry and psychology.

Big Data applications commonly comprise a range of various complex components. Data is extracted, collected, cleaned and transformed, stored and managed, analyzed, indexed and searched, as well as visualized. Accordingly, these applications touch many areas of computer science, such as database systems, text retrieval, data mining, machine learning and data visualization. For the scope of this paper, we thus define Big Data as the union over the range of tools and disciplines involved in collecting, storing, and analyzing large amounts of data originating from observing the interaction between users and devices (phones).

Indeed, independent of this paper's focus, the data in Big Data applications is frequently generated by logging user interaction. Examples range from recording queries from search engines to transactions from online shops. This data is commonly characterized by the three "v"s, its (i) velocity, (ii) volume, and (iii) variety. The former two indicate the speed at which data records arrive, and the large volume they amount to. The latter refers to the wide range of different data types and sources commonly involved in even a single application. Consider Google as a prime example for a Big Data use case, the collected data origins from a wide variety of services, such as its search engine, email service, and smartphone offerings (Android). Consequently, this data comes in a wide range of formats. Furthermore, already recording the queries issued to Google's search engine generates data tuples at a frightening rate, requiring extensive server farms.

Big Data applications commonly focus on data analysis. Researchers either attempt to (i) detect hitherto unknown trends and patterns, or (ii) "learn" new properties about known entities. The former discipline is commonly referred to as *data mining*, the latter as *machine learning*. One common application of data mining is the so-called market basket analysis, generating observations such as "customers who buy bread and butter commonly also buy beer". Machine learning in contrast attempts to detect hitherto unknown properties of objects. In classification for example, a common machine learning task, the software is shown a reasonably large training set of objects, for which the property of interest is known. By observing this *labeled* data, the software "learns" a computational model. Later, this model allows to classify other data items, i.e. automatically assign their labels. A credit card company could for example train a classifier by showing a large number transactions which have been labeled as either correct or fraudulent. In this example, features of interest could contain the country in which the card was issued, the one in which the transaction takes place, the type of shop and the amount involved. The resulting model could later be used to (dis-)approve transactions on a large scale.

Generating the computational models for data analysis is inherently labor intensive. In a trial-and-error fashion, data scientists repeatedly test the data for a certain hypothesis, fail, reconsider their assumptions and start over. In particular for machine learning, they have to (i) label a large training set, (ii) identify the most expressive data features to expose to the learning software, and (iii) extract these features from the raw data. Note, only the creation of the initial computational model is labor intensive. Later, this model can be used on a large scale, at little additional cost, thus benefiting from an economy of scale.

In contrast to what the title might suggest, Big Data is not so much characterized by the size of the accumulated data, but by the underlying motivation. Massive amounts of data are commonly collected *without an immediate business case*, but simply because it is affordable. This data, so it is hoped, will *later* answer questions, most of which yet have to arise. In addition to not knowing the later application scenario at the time of data collection, scientist also commonly cannot anticipate whether or not the data will display patterns of interest or support a certain hypothesis. In a Big Data context, data analytics thus become inherently post-hoc.

The proposed architecture resembles those employed in many other Big Data scenario. As depicted in Fig. 1, it gathers data from various *sources*, preferably without requiring user interaction. It forwards this data to a central data *storage*, where it is analyzed and mined for patterns, trends, and outliers. The resulting knowledge is then used in various *application* scenarios, such as academic research, therapy, or entirely new products.

Any piece of machinery allowing user interaction constitutes a potential data source. The user's *desktop* computer for example allows documenting the number of typographic errors, a potential symptom of stress and tiredness. Yet, modern IT has penetrated our everyday lives to an even higher degree. From coffee machines to cars and fridges, all gadgets in this world of *ubiquitous computing* (will) have some computational power, as well as the ability to communicate over the Internet. They can opaquely track the number of cups we require to get started in the morning, the frequency of out-of-schedule snacks, as well as our propensity to tailgate on the highway. Changes in any such parameter, so we claim, might indicate a shift in e.g., stress or aggression.

The introduction of smart watches and eyeglasses has further enhanced the ability to track human behavior. Additionally, recent advances in miniaturization have made *wearable sensors* small and affordable enough for everyday use. Only a few years ago, wrist-mounted accelerometers were the size of a bar of soap, and cost more than a desktop PC. Nowadays, the same technology fits to the size of a USB-stick, and sells for as little as 50$US. While it is thus possible to track the user's movements in the real world, we must also track online behavior. The average user spends a significant amount of our lives interacting with *Web applications*, such as social networks, shopping sites or online games. Most of these services offer a programming-interface which straightforwardly allows recording with who we communicate, for how long, what we buy, at which time, and how much we enjoy bashing orcs with magic swords in computer games. Any such behavior might correlate with traits of personality, while reported changes therein might aid a clinical practitioner in treating a psychiatric disorder.

The currently foremost source of behavioral data represents *Smartphones,* featuring a constant broadband connection, and exceeding the computational power of an early Pentium PC. They are intimate devices, carried on the body throughout the day. We commonly interact with our phone within the last half an hour before falling asleep, and no later than 30 min after waking up. (We will report the first exact data on this topic in the near future). The phones provide our most central platform to communicate, professionally, as well as to friends and spouses. The operating system allows tracking any phone-call or SMS, as well as any interaction with the installed software, the so-called apps. In addition, smartphones feature an entire set of sensors, such as for GPS signals and acceleration. It is thus possible to track the user's macro movement throughout the day, as well as e.g., her phones.

On the central server, the data is subject to various levels of analysis. These can range from simple counting (the number of times someone flicks on his phone), to a complex form of data mining and machine learning, especially for temporal patterns. Computer scientists have generated an entire range of methods for data analysis that yet have to be broadly applied to behavioral data. While far from trivial, one can essentially detect all patterns that are also apparent to a (trained) human eye. Riding on a train next to a passenger who keeps checking his phone, one might deduce that he is nervous, or bored. The human eye might come to the conclusion of the passenger to be nervous instead of being bored if the human brain then connects the observation of constant phone checking with the trembling of hands. The above software architecture might draw similar conclusions, albeit on the scale of thousands of phone users. Frequently, modern machine learning even outperforms the observational powers of humans. Duhigg reports the prominent case of a concerned father complaining to an American retailer for marketing pregnancy related products to his teen-age daughter [19]. The company's marketing department had indeed analyzed user's shopping behavior, trying to identify women in the second trimester of their pregnancy, who would

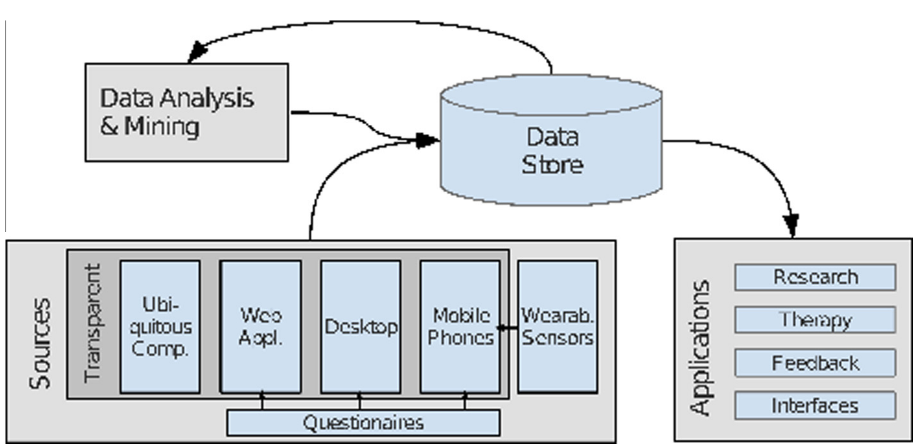

Fig. 1. A Big Data architecture in the context of psychometrics.

soon make baby-related purchases. As it turned out, the girl was indeed pregnant. The retailer's data-analytics team had outperformed the observational skills of her own father, who she was living with.

For the academic community, the above methodology yields an entire range of benefits. First, it generates behavioral data of *unprecedented quantity*. A single cellphone may well produce a thousand data points per day. Combined with the methods of data mining and machine learning, the approach thus enters an uncharted area of behavioral patterns and trends. Second, it is entirely *opaque* to the user, and does not require any explicit interaction. It can thus be employed to track user behavior for months and years, without overburdening the participant. Third, the approach avoids the dominant sources of bias in studying human behavior. In particular, it does not remind the user that he is participating in a clinical trial. Most notably, collected data measures the objective behavior (i.e., how he actually interacted with the phone), in contrast to subjective self-reports, which commonly suffer from a distorted or faulty recollection. Finally, data is transferred and analyzed *automatically*, rendering the proposed approach significantly cheaper than traditional paper-based methods. In summary, the approach thus increases data quantity and quality, while simultaneously unburdening participants as well as researchers.

Any conclusion automatically deducted from observations of human behavior can naturally be off. Typographic errors for example may as well be due to stress, as to a four-year-old yanking our sleeve. Thus, we do not propose to automatically generate diagnoses, but to quantitatively assist the medical practitioner, akin to a complete blood count. Following a multi-modal approach, simultaneously taking several different sensors into account, can further minimize error rates. For example, the combination of an increasingly monotonous voice, and progressing social withdrawal might indicate the worsening of a case of depression. Commonly, it will not be possible to determine absolute thresholds. Instead, one will rather investigate intra-personal change over time. For example, it will not be possible to deduce stress from the rate of typos exceeding x per 100 lines of written text. In contrast, a 50% increase in typos over 6 months could imply a substantial change in the patient's condition. Research will thus focus on detecting *changes over time*, the first statistical derivation of the observed signal. Likewise, it will frequently not be possible to detect trends applicable to an entire population of users. Yet, it will be feasible to detect *typologies*, i.e. classes of people exhibiting similar behavior. Data mining, in form of clustering and classification, can help find these typologies as well as the characteristic patterns in behavior.

Current research hypotheses in psychiatry and psychology

In two current studies, we monitor smartphones to track (i) the severity and course of depression as well as (ii) conspicuous usage of the Internet and phone. While these studies are decidedly small-scale, at least compared to the above technological vision, they are primarily intended to evaluate the validity as well as practicability of the proposed methodology.

Monitoring depression

In a first study, we currently employ smartphones to monitor depression. In particular we observe (i) app usage, (ii) social interaction, and (iii) macro movement of patients. In the context of depression, our central thesis is that these (and eventually most other) signals show less energy and dynamics, reflecting anhedonia and social withdrawal as central symptoms. To collect the necessary usage data, we have developed an app running on Android phones (version 4.0 and up), comprised of a software module for each of the three parameters, and one for encrypted data transfer to the server.

For app usage, we introduce the concept of an *app-session*, the time interval during which the user interacts with any one particular app. This implies that the app runs in the foreground, *and* the patient actively interacts with the phone. A session ends, when the patient (i) switches to another app, (ii) closes the app, or (iii) stops interacting with the phone. We thus record data tuples of the format: [app name, start-time, end-time]. In this context, we treat both, the lock-screen as well as the home menu, as just two additional apps. We thus also record, how often the user flicks the phone on (without actually unlocking it), and how much time he spends in the main menu.

Regarding *communication patterns*, we log in- and outgoing SMS as well as phone calls. These measures allow to infer the size and usage of the users social network. In this context, we only document with how many contacts the user regularly interacts, and who initiates the communication. The actual content of the calls (or SMS) is of no concern. Also, we are not interested in the actual identity of the communication partner. Hence, we anonymize their phone numbers, using cryptographic hash functions (SHA-512). Akin to app-session, we record time stamped data tuples, namely for calls: [anonymized number, start-time, end-time, in/out], for outgoing SMS: [anon. number, length in characters, time-sent], and for incoming SMS: [anon. number, length in characters, time-received, and time-read]. Additionally, we monitor the phone's address book, to estimate the total number of contacts, their changes over time, and the fraction, which the user actively communicates with.

For *macro movement*, we regularly record the user's geographic position. In the context of severe depression, we are commonly concerned whether a user even just leaves the apartment. Every 20 min, we thus estimate his location using GPS, recording the data tuple [time, latitude, longitude, accuracy]. Location estimation is additionally supported through available Wi-Fi signals and triangulation of cell-phone towers. Such functionality is readily provided by the phone's operating system. The resulting accuracy can be expected to range between 10 and 100 m, enough for macro movement, but not sufficient to track movement inside a building.

We are currently embarking on a study, observing patients suffering from major depressive disorder (MDD) over a period of four months. These patients are either treated with electro convulsive therapy (ECT), magnetic seizure therapy (MST) or deep brain stimulation (DBS) in the Department of Psychiatry and Psychotherapy at the University Hospital of Bonn. In order to establish a baseline, we install the app 1 month prior to the treatment, and continue measuring over the following 3 months of treatment. For the assessment of the current depressive symptomatology, we apply the Montgomery-Åsberg Depression Rating Scale (MADRS) as well as the Beck Depression Inventory (BDI-II) on a biweekly basis, as clinician- and self-rated questionnaires respectively. Due to drastic improvements commonly caused by the employed treatment methods, we expect to observe a significant signal-to-noise ratio. Furthermore we hypothesize that there is a significant correlation between the data concerning app usage, social interaction, macro movement and the applied questionnaires. We moreover hypothesize that this new method for monitoring the severity and course of depression is more sensitive to change, i.e. improvement of the depressive symptomatology (increase in activity patterns, resumption of social contacts) can be detected earlier.

Internet addiction & online social networks

Our second current study addresses a topic inherently linked to smartphone usage – Internet Addiction. As it does not constitute a distinct disorder in the DSM-V, but can only be found in the

Appendix, further research efforts are required to better characterize this emerging threat to our mental health. Until now, the diagnosis of Internet Addiction heavily relies on questionnaires and to a lesser extent on structural interviews. Given the mentioned shortcomings of both methods, the detection of behavioral addictions will profit enormously from the collaboration with computer science. Indeed, illustrating the validity and practicability of Psycho-Informatics represents one of the most important aims of this our research endeavor.

In a pilot study, we currently monitor the mobile phone behavior of $N = 100$ healthy participants for a duration of 6 weeks. This data thus provides the first reliable longitudinal statistics on diverse facets of smartphone usage and (mobile) Internet Addiction. We are particularly interested in fluctuations in the mobile phone behavior across the 6 weeks, but also search for stable – non-changing – facets. Questions to be answered are: How often is a phone exactly used each day? What are the most commonly used applications? When does a person go online for the first time in the morning? How often do they check their phones, mails or news? All these questions are impossible to grasp accurately by just asking participants. Instead, behavior needs to be recorded on the device. Indeed, a central point to the current study is the vast discrepancy between questionnaire based self-reports and actual phone usage. Additionally, we try to better characterize social activity. In this context, we explicitly do not limit ourselves to online social networks (such as Facebook), but study the size of the actual active social network reflected by phone-calls and SMS. Since we also record brain scans (structural MRI and resting state fMRI) and genetic material from participants, we are then able to correlate biological markers with social activity. Ideally, we would find a correlation between the active social network and certain areas of the brain. A recent study addressed a similar question, correlating amygdala volumes with the size of the social network [20]. This work however had to rely on self-reports, whereas we are able to measure the actual social activity, and in much finer detail. Ultimately, correlating Big Data with neuroscientific measures should carry far enough to eventually establish its own research direction of Psychoneuroinformatics.

Ethical aspects and data privacy issues of 'Big Data' research

The use of Big Data in research and therapy necessarily raises ethical concerns. Bordering mass surveillance, it realizes the vision of a "Gläserner Mensch", a transparent human. Data privacy thus takes on a central role, and the potential of abuse cannot be overestimated. While monitoring depression in a medical scenario fulfills the highest ethical standards, it could equally well be misused by an employer to secretly monitor his staff, or by an insurance company to reject at-risk applicants. This research is however not aimed at Digital Taylorism, a strategy that would surely backfire, but preventing, detecting and curing psychological disorders. Both, medicine and psychology, have worked on the vision of a transparent human since their very beginning. And for the same time-frame, both have had to handle sensitive data. There thus exists a proud tradition of confidentiality, whose methods can serve as blueprints for the deployment of Big Data technologies. Scaling and extending these concepts to an entirely new dimension is no mean feat, and will generate a significant amount of work for researchers, practitioners and occupational bodies.

From a different perspective, privacy concerns constitute somewhat of a side effect. Their denial would be entirely unethical. But, as this paper should outline, Big Data holds the potential to facilitate treatment of mental diseases. Hence, it would be equally unethical to bluntly deny its usage due to privacy concerns. Rather, the medical sciences have to follow another of their proud traditions: balancing risks and benefits on a case-by-case basis.

On a practical level, we currently follow a simple *two level privacy model*. We only collect usage data (i.e., behavior) that commonly needs to be aggregated to make any significant statement. In contrast, we refrain from collecting one-to-one written conversation, audio recordings, or video captures. Where applicable, text, audio and video are analyzed locally on the phone, and only the resulting markers and numeric values are sent to the server. In a current prototype for example, we estimate the size of the vocabulary from messenger applications on the phone, and only report resulting figures to the server.

On a different note, this paper does *not* propose an automatic diagnosis of psychiatric disorders. Instead, it suggests to assist the medical practitioner with additional information, in order to monitor the course of a disease and treatment. In certain cases, it may point a clinician in a certain direction while making an initial diagnosis. The ultimate responsibility, for false positives as well as negatives, remains with the clinician, as it has for two-thousand years. Whether one may be able to fully automatically diagnose mental disorders at some point in the future, we dare not hypothesize. Neither case would render the clinician obsolete, as communicating the findings to a patient, and even whether to do so at all, is a complex matter, not to be left to machinery.Proposed methodology could potentially serve as an early warning sign to a medical practitioner indicating that a patient might suffer from a condition. In this spirit, the paper compared the proposed methodology to a complete blood count. Neither technique makes a diagnosis, but assists a diagnosis, made by a clinical practitioner. The ultimate responsibility, for false positives as well as negatives, thus remain with the clinician, as it has for two thousand years. By providing additional quantitative information, we would hope to reduce the inherent actalepsy.

Conclusions and vision for the future

This paper introduces Psycho-Informatics, the application of Big Data to psychology and psychiatry. Highly sensitive, the suggested method collects, stores, and analyzes massive amounts of indicative data at little cost and without risks or stress for patients or participants. The paper outlines the technical vision, sketches the signals that can be detected, and illustrates the tremendous benefits over traditional methods of psychometrics. In particular, it suggests tracking user behavior with smartphones, a particularly rich and intimate source of data. This approach underlies two current studies, in the context of (i) depression and (ii) (excessive) usage of smartphones. The proposed methodology outperforms traditional methods in both quality as well as quantity. Namely, it avoids biased self-reports, and avoids altering the user's behavior, as the data is completely collected in the background. Furthermore, it collects data on a much finer granularity than conventional questionnaires and enables the search for temporal activity patterns. Additionally, there is no need to collect the data manually, as it is directly available in electronic form. While there are strong ethical concerns, these must not be allowed to evolve into thought-terminating clichés. Instead, they are to be addressed on a detailed level, case-by-case, following a rich tradition in medicine as well as psychology. These concerns being addressed appropriately, Big Data is about to revolutionize both psycho-sciences in research as well as therapy.

In the near future, the researchers will embark on numerous projects incorporating simple aspects of human-machine-interaction and wearable sensors. Already in the medium future however, focus will shift towards data analytics. Once the low hanging fruit has been picked (e. g. 'simple' descriptive statistical data on what is done how often on a smartphone), scientists will need to dig deeper inside data. Simple aggregate functions (e.g., count, or sum) will no longer suffice, but be replaced by mining for complex temporal patterns. Eventually, the entire range of methodologies from data mining and machine learning will have to be adapted to behavioral data. The effects on psychology and psychiatry will equal those of the massive change that the life sciences have undergone, and even fundamental research methodologies will have to be revisited. We will frequently hear of Psycho-Informatics and its sub-areas, such as Psycho-Neuro-Informatics. Most importantly, it will be possible, and not uncommon, to make an academic career, by solely studying data. Frequently, we will no longer design studies, but subject existing data to deeper analysis. Some of these datasets may have been conceived as a by-product of entirely different (non-academic) applications. In this context, scientists will have to learn to yield control. Research will shift from carefully constructed experiments on small parts of the population in a controlled environment, to massive longitudinal recorded data on tremendously large populations, full of errors and noise. Yet, so we hypothesize, signals will contrast from noise clearer than ever, due to the sheer amount of data.

To the same extend, Big Data will affect the daily routines of patients and clinical practitioners alike. The former will collect seemingly unrelated data to share with coaches and therapists. Provided with the necessary toolkits and expertise, the latter will be able to observe the course (and origins) of a disorder as well as the progress of treatment. This painted picture will be more accurate compared to previous self-reports, and of such fine granularity as to allow a highly individualized medication. Compared to traditional methods, this data-driven therapy will be cheaper, and consume less time from both, therapists as well as patients. Such technology requires tremendous research efforts. Already the establishment of meaningful metrics (as opposed to raw data tuples) as well as visual data exploration tools will be a laborious and ongoing effort. However, in the medium to long term, this data-driven therapy will become cheaper than traditional methods, consuming less time from both, therapists as well as patients. Most importantly, the approach extends the benefits of psychiatry and psychology far beyond treatment, into systems for early warning and mental wellness.

Conflicts of interest statement

None of the authors' reports a conflict related to the work described. The software mentioned is currently developed for research purposes only, no commercial exploitation of it is planned at this stage.

Acknowledgement

This work was partially funded in part by a grant awarded to C.M. by the DFG (MO-2363/2-1) and an independent investigator grant for the assessment of effects of deep brain stimulation for treatment resistant depression by Medtronic Inc. to TS.

References

[1] Costa e Silva JA. Personalized medicine in psychiatry: new technologies and approaches. Metabolism 2013;62(Suppl. 1):S40–4.
[2] Katschnig H. Quality of life in mental disorders: challenges for research and clinical practice. World Psychiatry 2006;5(3):139–45.
[3] Hamilton M. Rating scale for depression. J Neurol Neurosurg Psychiatry 1960;23:56–61.
[4] Hamilton M. HAMA Hamilton Anxiety Scale. In: Guy W, editor. ECDEU assessment manual for psychopharmacology, 193–198. Rockville, Maryland: National Institute of Mental Health; 1976.
[5] Montgomery S, Åsberg M. A new depression scale designed to be sensitive to change. Br J Psychiatry 1979;134:382–9.
[6] Schlaepfer TE et al. The hidden third: improving outcome in treatment-resistant depression. J Psychopharmacol 2012;26(5):587–602.
[7] Della Pasqua O, Santen GW, Danhof M. The missing link between clinical endpoints and drug targets in depression. Trends Pharmacol Sci 2010;31(4):144–52.
[8] Sadeh A. The role and validity of actigraphy in sleep medicine: an update. Sleep Med Rev 2011;15(4):259–67.
[9] Sadeh A, Acebo C. The role of actigraphy in sleep medicine. Sleep Med Rev 2002;6(2):113–24.
[10] Ancoli-Israel S et al. The role of actigraphy in the study of sleep and circadian rhythms. Sleep 2003;26(3):342–92.
[11] Stanley N. Actigraphy in human psychopharmacology: a review. Hum Psychopharmacol 2003;18(1):39–49.
[12] Young K. Internet addiction: the emergence of a new clinical disorder. Cyberpsychol Behav 2009;1(3):237–44.
[13] Bianchi A, Phillips J. Psychological predictors of problem mobile phone use. Cyberpsychol Behav 2005;8(1):39–51.
[14] Shaw M, Black DW. Internet addiction: definition, assessment, epidemiology and clinical management. CNS Drugs 2008;22(5):353–65.
[15] Montag C, Jurkiewicz M, Reuter M. Low self-directedness is a better predictor for problematic internet use than high neuroticism. Comput Hum Behav 2010;26(6):1531–5.
[16] Ko CH et al. The association between Internet addiction and psychiatric disorder: a review of the literature. Eur Psychiatry 2012;27(1):1–8.
[17] Kosinski M, Stillwell D, Graepel T. Private traits and attributes are predictable from digital records of human behavior. Proc Natl Acad Sci USA 2013;110(15):5802–5.
[18] McCrae RR, John OP. An introduction to the five-factor model and its applications. J Pers 1992;60(2):175–215.
[19] Duhigg C. How companies learn your secrets. The New York Times; 2012.
[20] Bickart KC et al. Amygdala volume and social network size in humans. Nat Neurosci 2011;14(2):163–4.