Scaling-out Longitudinal Clinical Analytics with Dataflow Processing

Joan Saez Pons, Cosmin Stamate, David Weston, George Roussos
Department of Computer Science and Information Systems
Birkbeck College, University of London
Malet Street, London, U.K.
{joan, cosmin, dweston, g.roussos}@dcs.bbk.ac.uk

***Abstract*—There are two key ingredients in supporting high-frequency and continuous clinical assessment of patient populations at scale: first, the availability of validated metrics of disease progression which reliably capture the longitudinal variations of symptoms; and second, the ability to compute these metrics on the fly over multiple concurrent streams of sensor data captured at home or in the community. In this paper, we describe the design, development and validation of PDkit, a comprehensive data science toolkit for Parkinson's Disease, and explore the dataflow paradigm as a means to provide salable performance. Our aim is to contribute towards the development of robust clinical outcome measures for therapeutic trials and to support longitudinal investigations of disease mechanism through the analysis of data collected from wearables and smartphones. The PDkit is released as open source and offers a succinct interface for interactive collaborative data exploration. Moreover, it enables the composition of data processing pipelines for tremor, tapping, bradykinesia and gait tests with the view to support horizontal scalability over common Cloud infrastructures on production workloads. Specifically, we report on our early experiments executing PDkit pipelines using Apache Beam, a unified dataflow multi-runtime stream processing engine. Our long-term aim is to provide the PD research community with the tools needed to individually tailor treatment plans and to empower patients to become more involved in their own care.**

***Index Terms*—Parkinson's Disease, clinical assessment, dataflow processing, sensor data streaming.**

I. Introduction

Rapidly ageing global populations bring about accelerated growth in the prevalence of long-term neurodegenerative diseases including Huntingtons, Parkinsons and Alzheimers disease and other dementias. Such diseases progressively affect the neurones of the human brain leading to debilitating conditions and are incurable. Parkinson's Disease (PD) in particular is associated with a wide spectrum of symptoms including tremor, slowness of movement and freezing, swallowing difficulty, sleep-related difficulties and psychosis [7]. Since there is no cure, symptom management is a life-long process and includes pharmacological treatment typically with L-Dopa, physiotherapy, and surgery in its latter stages [25].

Regular monitoring of People with Parkinsons (PwPs) and adjustment of medication is a key ingredient of typical clinical care pathways. However, with over 10 million PwPs worldwide and a rapidly increasing patient population , current approaches to monitoring and care are becoming overwhelmingly unaffordable and do not scale. One way to confront this challenge is offered by the wider availability of smartphone apps and wearables employed to monitor symptoms is bringing about a fundamental transformation in the way PwPs can be assessed: The application of these technologies enables the unsupervised, and at high-frequency or continuous measurement of motor and non-motor performance of a large patient population [3], [11]. Mirroring patterns of contemporary data production in other domains, this paradigm shift in the clinical assessment of PD leads to the tremendous increase in the availability of patient performance data. In this setting, manual analysis of data is no longer viable. Instead, it is imperative to adopt a software-based approach so that outputs of clinical relevance can be computed automatically and presented to researchers, clinicians and patients in an intuitive manner [4].

To this end, in this paper we present the design, development and validation of a comprehensive software toolkit for the management and processing of patient data captured continuously by wearables [5], [13] or by high-use-frequency smartphone apps [1], [10], [15]. The toolkit facilitates the application of a data science methodology to the analysis of this information incorporating an extensive collection of methods and techniques selected from the PD literature. Although inherently flexible, in the development of PDkit we have prioritized functionality critical to therapeutic clinical trial delivery rather than general patient care.

Presented in Sections III and IV below, the PDkit is released as open source under the permissive MIT License as we believe that open and inclusive access to its features will provide a key ingredient towards realizing the promise of mobile and wearable technology for PD. Further, we report on our experimentation using patient data collected through our own work and using open data sets, which suggest that achieving the above goals requires the application of longitudinal metrics rather the one-shot approach commonly in practice today. We discuss this alternative approach in Section V where we demonstrate how it can work in practice.

Working at full population scale to conduct longitudinal assessments demands that systems infrastructure must scale out. In Section VI, we demonstrate how the pipeline-base approach adopted for the development of the PDkit facilitates modern stream processing architectures supporting scalability on modern cloud platforms. We conclude with a summary of our finding and directions for future work in Section VII.

978-1-5386-9151-9/19/$31.00 ©2019 IEEE

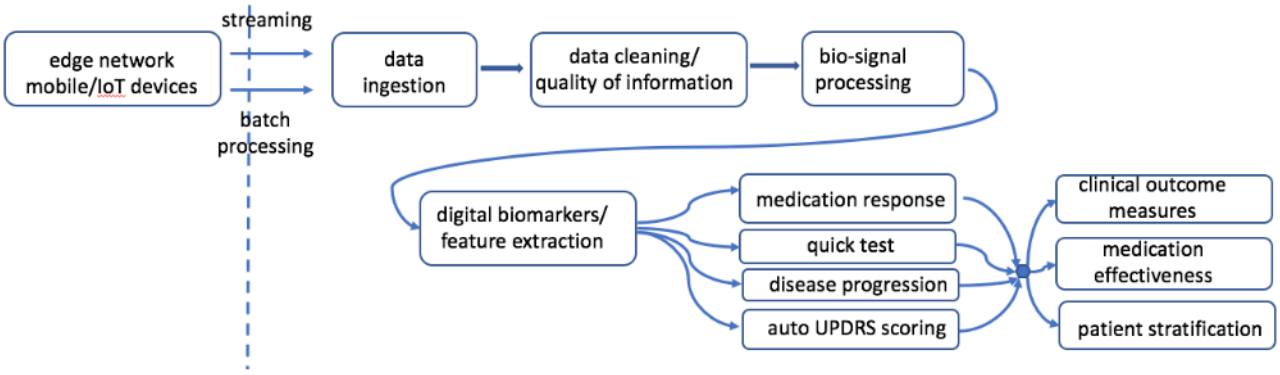

Fig. 1. The PDkit processing pipeline for high-frequency and continuous measurement of motor symptoms for Parkinson's Disease.

II. Background and Related Work

Over the past decade, several smartphone apps and wearable systems [1], [6], [13], [14], [27], [33] have been developed to address the needs of PD: the mPower app developed for iOS by Apple and Sage Bionetworks in the US; the uMotif app developed with NHS SBRI Healthcare funding; the Wearable Companion app developed by the MJ Fox Foundation and Intel; the mHP app for Parkinsons developed by myHealthPal; the Verily app in collaboration with ParkinsonNet in Holland; and several others. We refer to these apps collectively as following a High-Frequency pattern of assessment in contrast to standard clinical practice which is carried out relatively infrequently. This terminology reflects the fact that the apps are able to carry out measurements of most elements of motor and cognitive performance of PD patients multiple times per day or even continuously when wearables are used in addition to a smartphone [9].

In our own work [17], we developed cloudUPDRS, the first smartphone app to achieve certification as a Class I medical device by the Medicines & Healthcare products Regulatory Agency for the clinical assessment of the motor symptoms of Parkinson's. cloudUPDRS is based on the Unified Parkinson's Disease Rating Scale [5] and the PDQ39 questionnaire [29], and incorporates cloud-based Big Data management and analytics microservices to generate objective and precise assessments of motor performance [31]. Following Part III of MDS-UPDRS, patients use the app at home to record high-precision sensor measurements while performing a series of simple actions with each limb, such as tapping the screen to assess bradykinesia and holding the phone on their knee to assess tremor [10]. The data captured by the app is then used to calculate the clinical UPDRS score through the application of a biomedical signal processing pipeline (cf. Figure 1). Additional longitudinal analytics can be performed subsequently to enable trend analysis and patient stratification [28].

III. PDkit: Features and Functionality

As noted in Section II, there is intense interest in the development of automated assessments for PD symptoms through the use of smartphone apps and wearables. However, each system proposed in the literature adopts one or more performance metrics from a diverse variety of alternatives proposed in the literature. As such, it is often difficult to directly compare the performance of alternative approaches or indeed compare novel techniques against a state-of-the-art baseline. To address this, we have developed PDkit as an open source toolbox that supports such experimentation through the development of a wide variety of standard performance metrics for PD supporting a simple to use set of abstractions. The toolkit is structured around the PD data processing pipeline presented in Figure I.

The PDkit can play a critical role supporting therapeutic development and cost-effective clinical trial evidence collection, by facilitating the development of: (i) detailed clinical outcome measures that enable for example the early identification of problems such as medication side-effects, (ii) robust quantitative metrics of disease progression computed automatically from the data, (iii) individualized patient profiles leading to personalized assessment, and (iv) patient stratification through longitudinal analytics. In the following Sections we briefly outline the four areas of functionality implemented.

A. Quality of Information

The first stage of the PDkit pipeline supports the preparation of raw data for analysis, a process often also referred to as data wrangling, providing methods and techniques for assessing, structuring, cleaning, and rolling up source data. This is typically a time-consuming and unattractive aspect of data science but essential to obtain accuracy, precision, consistency, and completeness. PDkit implements a variety of such methods ranging from standard integrity tests to identify missing data, faulty sensing elements in the collecting device and corrupted data items, to advanced techniques for the verification of the correctness of the data collection process using deep learning [16], [17].

B. Bio-Signal Processing

The second stage of the pipeline provides signal processing techniques for the computation of over 200 distinct biome-

chanical metrics characterizing motor and non-motor performance. For example, for tremor measurements, the application of the Fourier and wavelet transforms offer metrics such as dominant tremor frequency and power, and also facilitate filtering of non-essential frequencies captured in the signal [9]. For bradykinesia measurements, signal processing calculations relate to calculating the kinesia, akinesia, dysmetria and arrhythmia scores. For gait measurements, signal processing calculations estimate stride length and frequency, symmetry of walking, lateral swinging, magnitude of acceleration, turning performance (such as time and number of steps required) as well as detection of freezing. In the future, the toolkit will provide signal processing techniques for voice data to extract dysphonia metrics and processing speed for Stroup tests.

C. Digital Biomarkers

The third stage of the pipeline employs the metrics calculated by the bio-signal processing module to compute unitary and composite digital biomarkers. This process is akin to the process of feature engineering in data science tailored to the PD setting whereby metrics derived from bio-signal processing calculations are used to identify or compose candidate features. Such features may relate to a single metric for example turn agility during walking [15] represented for instance by the number of steps required to complete a full turn, or can be applied on tuples combining several metrics. Moreover, digital biomarkers can be constructed either using a single motor performance observation from the raw data set taken at a particular point in time, or defined as the statistical distribution of the full sample taken over a specified period of time following a particular sampling strategy. PDkit implements a variety of realted digital biomarkers reported in the literature [24] as well as novel ones as they emerge for example selected proposals from the Parkinsons Disease Digital Biomarker DREAM Challenge organized by the Michael J. Fox Foundation. Closely related to the calculation of the digital biomarkers is the development of predictive models for estimate disease state or severity.

D. Clinical Rating Scale Calculation

FInally, PDkit provides calculations that map computed digital biomarkers to clinical rating scales notably those published by the Movement Disorders Society. Other rating scales will also be considered such as the Parkinson's Disease Composite Scale and the OTS score developed specifically for handheld devices at Uppsala University. Such mapping can be achieved through the implementation of clustering and classification algorithms associating biomarkers to rating scale scores [4]. In future versions of the PDkit we aim to incorporate general facilities to generate such mapping from a particular data set as well as pre-trained models [4] that can achieve this goal for a selected set of digital biomarkers depending on raw data availability.

IV. PDkit: Use Cases

The PDkit toolkit is available under the standard PyPI package management system for python and also full source code is available at `https://github.com/pdkit/` including Jupyer Notebooks demonstrating its use and extensive documentation `https://pdkit.readthedocs.io/` including bibliographic references for each of the features implemented.

To highlight some of the key use cases of the toolkit, in the section we include a description of a typical processing pipeline receiving raw tremor input from the accelerometer of a smart watch or mobile app and returning the UPDRS score associated with this test (the code below is slightly amended to facilitate presentation, full examples can be found on the github repo above).

```
>> tp = pdkit.TremorProcessor()
>> ts = pdkit.TremorTimeSeries()
>> amplitude, freq = tp.amplitude(ts, 'welch')

>> testResultSet = pdkit.TestResultSet()
>> testResultSet.process()

>> clinical_UPDRS =
     pdkit.Clinical_UPDRS(labels, testResultSet)
>> clinical_UPDRS.predict(measurement)
```

In the above source code example, an acceleration time series obtained by a wearable or a smartphone app is first processed to extract a measure of the associated tremor amplitude using the Welch method. Subsequently, multiple assessments are aggregated to train a classifier which can map a new tremor observation to a clinical score on the UPDRS scale. The initial data cleaning of the incoming raw data is integrated transparently into the TremorProcessor class.

In the example below, we replicate the first stage of the above pipeline for walking tests. In this case gait measurements using an acceleration sensor carried by the patient at the pelvis, is employed to extract standard performance metrics such as regularity and symmetry of walking:

```
>> ts = pdkit.GaitTimeSeries().load(filename)
>> gp = pdkit.GaitProcessor()
>> step_regularity, stride_symmetry =
         gp.walk_regularity_symmetry(ts)
```

Finally, the example below shows how raw measurements can be used to train a model for UPDRS scoring in the absence of validated clinical assessments. In this case, a clustering approach has been adopted to discriminate between the different levels of the clinical rating scale.

```
>>updrs = pdkit.UPDRS(testResultSet)
>>updrs.score(observation)
```

V. Longitudinal Analytics

Recent advances in the pathophysiology of Parkinsons disease, including genetic and biochemical causes, have considerably expanded the understanding of its pathogenic processes and as well as of pharmacological responses to therapeutic interventions. Despite this progress, the development of clinical biomarkers for the evaluation of disease progression remains highly challenging [30]. This is as much due to individual patients having markedly different symptom constellations, progression rates, and treatment responses, as due to the dramatic day-to-day variation of the symptoms of a single

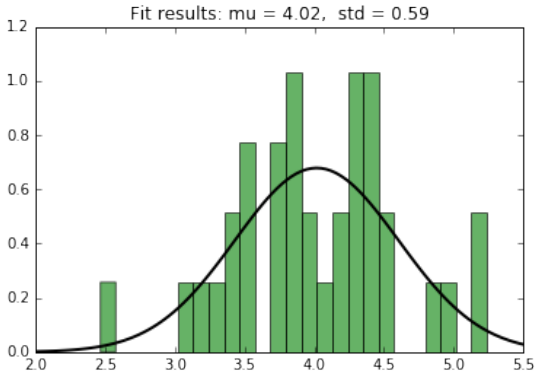

Fig. 2. Longitudinal-composite disease progression metric (April 2016).

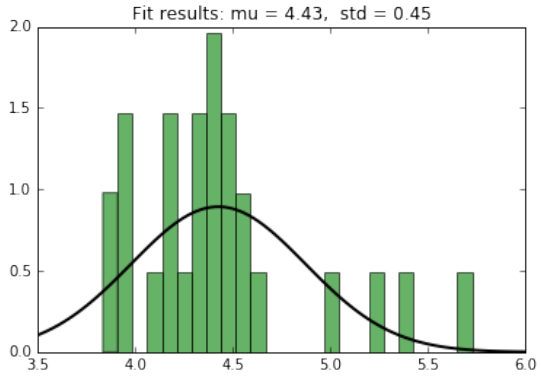

Fig. 3. Longitudinal-composite disease progression metric (July 2016).

patient. Current measures of disease progression commonly used in clinical practice, such as the UPDRS [5] and PDQ39 [29], do not adequately account for this variation.

Furthermore, such clinical measures have several limitations [26]: First, their application is laborious and costly because they require the direct involvement of a member of the clinical team and as a consequence can only be carried out infrequently. Second, although clinical measure protocols are detailed and formally structured they are nevertheless carried out at relatively coarse-grain granularity without the use of specialised measurement instrumentation. Finally, despite generally good internal consistency in the application of these measures, they still depend on subjective estimations of patient performance by the clinician. Collectively, these constraints suggest that a single calculation of a particular clinical measure at a specific point in time does not accurately describe the actual state of the patients condition, even when conducted under controlled conditions in a clinical setting.

In this paper, we argue that such unit assessments are better understood as drawing a single observation from a highly dynamic process depicting the clinical presentation of PD. From this viewpoint, attempting to characterise the overall pattern of PD presentation from a single observation is clearly not viable. Instead, we suggest that a more effective approach is to consider the problem of developing disease progression indicators by following a process sampling paradigm. According to this approach, indicators are based on the statistical properties of the full sample of symptom measurements obtained over a defined sampling time-frame and following a suitably tailored sampling strategy.

Adopting this approach at scale has not been feasible until now due to the very high cost associated with conducting frequent assessments by specialist clinical staff [32]. Wearable and mobile technologies however enable unsupervised monitoring of a large patient population by conducting precise and objective measurements of motor and non-motor performance multiple times per day at almost no cost. It is now thus possible to implement a practical and comprehensive sampling strategy for PD executed independently by patients at home.

Yet, translating these measurements into a marker of disease progression remains challenging. To this end, we have proposed a novel PD progression indicator adopting the process sampling approach.The goal of this approach is to establish a formal methodology for the reconstruction of the underlying statistical distribution of PD presentation from smartphone data and design a novel disease progression indicator derived from it. Our goal is to establish this indicator as a practical, reliable, precise and objective alternative that can moreover accurately determine the efficacy of therapeutic agents.

Our initial experiments with high-frequency smartphone technology provides strong support for the observation that patient motor performance exhibits complex patterns of variation. Figure 2 shows symptom measurements for a single patient over a three-week period of monitoring. The specific digital biomarker displayed in this case is the strength of the dominant tremor frequency for right leg agility measurements (we note that similar patterns of variation are observed for all digital biomarkers associated with affected limbs). Figure 2 also provides annotations showing the corresponding UPDRS Part III test score calculated by cloudUPDRS. Clearly, no regular patterns emerge during this period despite the fact that the patient follows a predictable medication routine (administered 4 times per day). For example, note that for almost any 2-hour slot during, the corresponding score may range from 1 to 3 points on the UPDRS scale. Overall, considering this degree of variation in symptom presentation the data suggests limited value in assessing disease progression by comparing two UPDRS scores taken several weeks or indeed a few months apart as daily fluctuations are essentially indistinguishable from the underlying disease progression trend.

Further, we note that for a specific patient each individual test and its associated digital biomarkers have variable inferential power in establishing disease progression. This observation is corroborated by the literature which suggests that five to six clinically distinct factors typically suffice to provide high correlation to the patient's overall Part III UPDRS score. This should not be surprising considering that clinical measures explores exhaustively the full range of possible motor symptoms caused by PD, but a specific individual would typically present a much smaller number of symptoms which

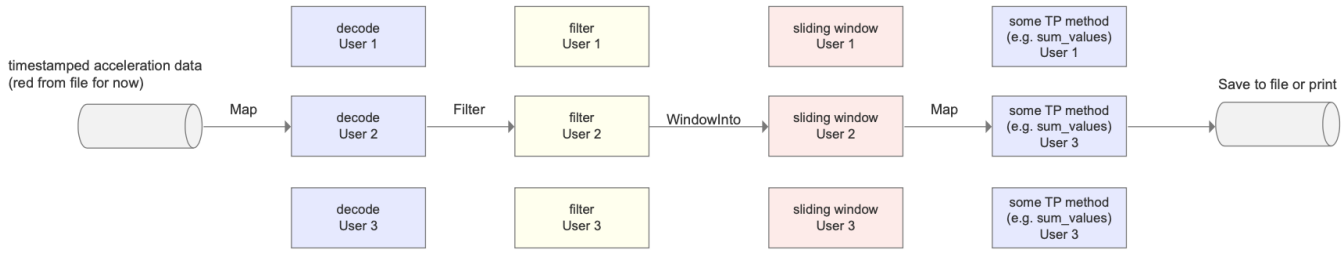

Fig. 4. Streaming PDkit processing pipeline using Apache Beam primitives.

dominate their overall score. In cloudUPDRS, we have developed a machine learning methodology which can successfully identify the appropriate subgroup of tests for a specific patient which offer the highest predictive power of their overall motor performance.

Overall, the combination of the high-frequency pattern of assessment with the bounded context for the interpretation of measurements enforced through user interaction design by the cloudUPDRS app, and the ability to identify the most critical clinically distinct factors represented by specific digital biomarkers for a particular individual, enables the effective statistical description of PD from the process sampling standpoint. To illustrate this point, in Figure 3 we aggregate all calculations of the same digital biomarker presented in Figure 2 taken during April 2016 (25 samples in total) into the histogram depicted in Figure 3, and those taken during July 2016 (43 samples in total) on the right. Both histograms suggest a normal distribution with distinct characteristics and we have further explored this suggestion beyond visual appearance. We summarise our main findings below.

First, the Shapiro-Wilk test was applied to both samples to establish whether they are normally distributed and confirm that this is indeed the case. Further, we establish that the two samples from April and July correspondingly are drawn from different distributions by applying the one-way analysis of variance (One-way ANOVA) and the Wilcoxon rank-sum tests. Both provide a strong indication that the two samples correspond to distinct distributions. Finally, we observe that the mean of each distribution could provide an indication of patient performance and we note that this would imply a deterioration of symptoms with a shift to the mean from 4.03 to 4.32 which would suggest a UPDRS score of 2 in both cases so it would be not possible to distinguish the change by using UPDRS only. We also note that the patient reported a noticeable drop in his motor performance during this period. Overall, in view of these observations it is reasonable to explore the hypothesis that a disease progression indicator developed according to the process sampling approach as described above can be effective.

Using the PDkit and a data collection app using the Vivosmart device, we are currently undertaking an exhaustive investigation of the statistical distribution of singleton and combinative digital biomarkers calculated from high-frequency measurements over different time-spans ranging from several days to a few weeks. To achieve this, we employ QoI techniques from the PDkit to pre-process the captured data so as to ensure high quality, and subsequently compute a full range of digital biomarkers. We anticipate that the resulting distribution is likely to follow a normal distribution as discussed above and to employ its defining features such as the mean and variance, as the basis for the generation of a disease progression indicator.

VI. Stream Processing

Supporting automated symptom assessment for PwPs at population scale requires the ability to process large amounts of data generated concurrently as an unordered and potentially unbounded stream. In this setting, data may also be received with delays that are not impossible to quantify precisely, particularly data are relayed via an unreliable edge system. Further, as noted in the previous Section, the calculation of longitudinal analytics requires aggregation operations which suggests the need to apply assessment metrics over multiple windows of the incoming data.

To address these requirements we are developing a streaming version of the PDkit, which in addition to matching the typical clinical workflow also facilitates scalability to full population scale. To provide transparent support across multiple cloud-based platforms while maintaining efficiency, the toolkit adopts a dataflow programming approach [21] as its main architectural pattern. This approach allows application developers to access toolkit functionality via a standard python interface irrespective of the underlying computing substrate, and thus reuse the same codebase unchanged. Dataflow systems achieve this by modelling a program as a directed graph of the data flowing between operations [23]. This approach generates efficient general purpose code that can be mapped automatically onto practically any modern cloud architecture [22] thus avoiding very specialized, non-reusable code which is typically required by alternative high-performance paradigms such as message-passing [21].

To implement the dataflow approach in PDkit we employ the Apache Beam system due to the fact that it offers distinct advantages: firstly, it provides a unified conceptual model for stream processing; second, Beam APIs are accessible via a

python SDK which we have used to extend PDkit through PD data-processing pipelines; and last nut not least, it provides a variety of so-called runners, that is transparent mappings to the underlying execution environment, for several popular distributed processing backends including Apache Flink and Spark and Goolge Cloud Dataflow.

The streaming version of PDkit is currently under active development with the first production release planned for January 2019. To demonstrate the natural mapping of the PDkit concepts into Beam pipelines, we have implemented a variety of typical processing scenarios. For example, the tremor assessment pipeline detailed in Section IV above can be easily be translated to the Beam API for multiple concurrent users as depicted in Figure 4: the process consists of a series of Beam `ParDo` operations for filtering, mapping and feature extraction, with the incoming acceleration data streams easily mapped between the PDkit primitive data types and the `PCollection` data objects employed by Beam.

VII. Discussion and Conclusions

In this paper we have considered composite metrics of disease progression in the context of Parkinson's Disease, which reliably capture the longitudinal variations of common motor symptoms. At the core of this approach is the observation that better measures can be developed by aggregating multiple symptom measurements carried out over a period of several days. Wjhile this approach is more effective in tracing symptoms is also presents a considerable user experience challenge. This suggests that a passive monitoring approach using wearables or a combination of passive and active elements would be preferable. Further, we explored dataflow streaming as the main enabler for the computation of these metrics on the fly over multiple concurrent streams of sensor data captured at home or in the community concluding that they complement well the sensor pipeline approach adopted for the PDkit.

Acknowledgments

This research is funded by the Michael J. Fox Foundation for Parkinson's Research (MJFF) with Grant ID 14781 for the project entitled "A Scalable Computational Data Science Toolbox for High-Frequency Assessment of PD" awarded under its Computational Science 2017 programme.

References

[1] S. Arora et. al. Detecting and monitoring the symptoms of Parkinsons disease using smartphones: A pilot study, Parkinsonism and Related Disorders, 21(6), 650653, 2015.

[2] C. Bettini et.al. A survey of context modelling and reasoning techniques, Pervasive and Mobile Computing, Volume 6, Issue 2, 161-180, 2010.

[3] B.M. Bot et. al. The mPower study, Parkinson disease mobile data collected using ResearchKit, Scientific Data, Mar 3, 2016.

[4] N.F. Fragopanagos et. al. Towards Longitudinal Data Analytics in Parkinsons Disease. Proc. 1st Int. Conf. on IoT and Big Data Technologies for HealthCare, June 15-16, Budapest, Hungary, 2016.

[5] C. G. Goetz, et al. Testing objective measures of motor impairment in early Parkinsons disease: feasibility study of an at-home testing device. Mov. Disord. 24, 551556,2009.

[6] N. Kostikis et. al. Towards remote evaluation of movement disorders via smartphones, Proc. IEEE Eng. Med. Biol. Soc., 52403, 2011.

[7] J. Jankovic, "Parkinsons disease: clinical features and diagnosis", J. Neurology Neurosurgery and Psychiatry, vol. 79, no. 4, pp. 368-76, 2008.

[8] A. Jha at. al. The cloudUPDRS smartphone app: home monitoring for Parkinsons Disease. The Future of Medicine Conference, Royal Society of Medicine, May 19, London, UK, 2016.

[9] P. Kassavetis et. al. Developing a tool for remote digital assessment of Parkinsons disease. Movement Disorders Clinical Practice, 3(1), 2015.

[10] S. Kueppers et. al. From Wellness to Medical Diagnostic apps: The Parkinsons Disease Case. Proc. Int. Conf. on Personal, Pervasive and mobile Health, June 14-16, Budapest, 2016.

[11] C.Y. Lee et. al. A Validation Study of a Smartphone-Based Finger Tapping Application for Quantitative Assessment of Bradykinesia in Parkinsons Disease, PLoS ONE 11(7): e0158852, 2016.

[12] R. Lemoyne et. al. Implementation of an iPhone for characterizing Parkinsons disease tremor through a wireless accelerometer application. Proc. IEEE Eng. Med. Biol. Soc., 49548, 2010.

[13] W. Maetzler et.al. Quantitative wearable sensors for objective assessment of Parkinsons disease, Movement Disorders, 28(12), 16281637, 2013.

[14] E.C. Neto et al. Personalized hypothesis tests for detecting medication response in Parkinson disease patients using iPhone Sensor data, Pac. Symp. Biocomput. 2016; 21: 273284.

[15] D. Pan, R. Dhall, A. Lieberman and D.B. Petitti. A Mobile Cloud-Based Parkinsons Disease Assessment System for Home-Based Monitoring, JMIR mHealth and uHealth, 3(1), e29, 2015.

[16] C. Stamate et. al., Deep Learning Parkinsons from Smartphone Data, Percom 2017, 2017.

[17] C. Stamate et. al., The cloudUPDRS app: A Medical Device for the Clinical Assessment of Parkinson's Disease, Pervasive and Mobile Computing, in press, 2018.

[18] S. Krishnan et. al., Interactive data cleaning for statistical modeling. VLDB9, pp. 948-959, 2016.

[19] R.C. Martin, Agile Software Development, Principles, Patterns, and Practices. Pearson, 2013.

[20] A. Kyrola et. al. GraphChi: large-scale graph computation on just a PC, OSDI12, pp. 31-46, 2012.

[21] Y. Low et. al. Distributed GraphLab: a framework for machine learning and data mining in the cloud, VLDB 5, pp. 716-727, 2012.

[22] I. Gog et. al. Musketeer: all for one, one for all in data processing systems, EuroSys '15. ACM, New York, NY, USA, Article 2, 16 pp., 2015.

[23] D. Yan et. al. Big Graph Analytics Systems, SIGMOD '16. ACM, pp. 2241-2243, 2016.

[24] M. Delenclos et. al. Biomarkers in Parkinson's disease: Advances and strategies, Parkinsonism Relat. Disord. 22 (Suppl 1): S106S110, 2016.

[25] A. H. V. Schapira, M. Emre, P. Jenner, W. Poewe, "Levodopa in the treatment of Parkinson's disease", European Journal of Neurology 16, pp. 982-989, 2009.

[26] M. Delenclos et. al. Biomarkers in Parkinson's disease: Advances and strategies, Parkinsonism Relat Disord. 2016 January ; 22 (Suppl 1): S106S110

[27] J.F. Daneault et. al. Using a smart phone as a standalone platform fordetection and monitoring of pathological tremors. Front Hum Neurosci, 6,357, 2008.

[28] N.F. Fragopanagos et. al. "Towards Longitudinal Data Analytics in Parkinson's Disease." Proc. 1st Int. Conf. on IoT and Big Data Technologies for HealthCare, June 15-16, Hungary, 2016.

[29] C. Jenkinson, R. Fitzpatrick, V. Peto, R. Greenhall and N. Hyma. The Parkinsons Disease Questionnaire (PDQ-39): development and validation of a Parkinsons disease summary index score. Age Ageing, 26(5), 353357,1997.

[30] DJM McGhee et al. A systematic review of biomarkers for disease progression in Parkinsons disease. BMC Neurology 13:35, 2013.

[31] S. Newman. Building Microservices: Designing Fine-Grained Systems.OReilly Media, 2015.

[32] National Institute for Health and Clinical Excellence. Parkinsons disease: diagnosis and management in primary and secondary care: National cost-impact report. NICE clinical guideline no. 35, 2006

[33] A. Rodriguez-Molinero et al. Validation of a portable device for mapping motor and gait disturbances in Parkinsons disease. JMIR Mhealth Uhealth,3(1), e9, 2015.[