# Dear applicant,

# This form is meant to support you in preparing your preproposal. When filling out the form, please remove this box *and* the explanations below the section headers (##).

# Guidelines proposal COMMIT2DATA – BIG DATA, HEALTH, & HUMAN PERFORMANCE

# This proposal should contain a sufficient description of a Big Data or Data Science Project that is feasible to be completed within one month. This includes descriptions of scientific, practical, and innovative value of the proposed project. The project should be conducted by a consortium of at least two scientific project partners, one public partner, and one end user. Federations and professional Clubs will be accepted as business partners. The findings of the projects need to be presented at an appropriate scientific conference.

Please submit your application to VIPC via Nestor before May 31th. To submit, a pdf format is required. In order to process the application data properly, the pdf file should not contain any security lock. This is the only electronic format that guarantees that the application will be received in exactly the same form as it has been sent.

Please complete the application in **English**. Do not exceed the stated maximum number of words for each item on the form and use a 10-point font size, except for references to the literature (APA-Style), which may be given in 9-point.

### BASIC DETAILS

**1a Title of the proposal**

*Early detection of symptoms of Parkinson’s Disease using finger movements during typing: a machine learning approach.*

### 1b Details of the main applicant

*## Please provide complete information as requested. The main applicant is responsible for the*

*## financial accounting for the part of the project that is funded by NWO.*

Name: Simon Coopmans

Affiliation:

Address of Knowledge Institute: Department for Human Movement Sciences, University Medical Center Groningen, University of Groningen, A. Deusinglaan 1 9713 AV Groningen

Preference for correspondence in English: **yes**/no

Email: s.c.coopmans@student.rug.nl

**1c Co-applicant(s)**

## Please provide complete information for the co-applicants as requested.

## Please fill in the details below for each co-applicant.

Name: Martijn van Haren

Affiliation:

Address of Knowledge Institute: Department for Human Movement Sciences, University Medical Center Groningen, University of Groningen, A. Deusinglaan 1 9713 AV Groningen

Preference for correspondence in English: **yes**/no

Email: m.van.haren@student.rug.nl

Name: Thijme Langelaar

Affiliation:

Address of Knowledge Institute: Department for Human Movement Sciences, University Medical Center Groningen, University of Groningen, A. Deusinglaan 1 9713 AV Groningen

Preference for correspondence in English: **yes**/no

Email: T.E.Langelaar@student.rug.nl

Name: Sanne Westerveld

Affiliation:

Address of Knowledge Institute: Department for Human Movement Sciences, University Medical Center Groningen, University of Groningen, A. Deusinglaan 1 9713 AV Groningen

Preference for correspondence in English: **yes**/no

Email: s.westerveld@student.rug.nl

**1d Composition consortium**

*##* *Fill in in the table below the involved persons, their affiliations and their representative research area (Data science, Life and Health sciences, Creative industry science, Social sciences).*

*##* *Make a clear distinction between the scientific researchers, (representatives of) private or public partners and end-users (citizens, patients, health professionals).*

*## Extra rows can be added or rows can be deleted*

|  |  |  |  |
| --- | --- | --- | --- |
| **Consortium** |  |  |  |
| ***Partner(s) research institute(s)*** | ***Institute/organisation*** | ***Role*** | ***Research area*** |
| Dr. M. Kempe | Department for Human Movement Sciences, University of Groningen. | PhD – Scientific support | Data Sceince |
| Dr. W.R. Adams | School of Computing & Mathematics, Charles Sturt University, N.S.W., Australia | Data provider | Life and Health sciences |
| **End user participation** |  |  |  |
| Patients suffering from Parkinson’s Disease |  | Data Provider- End user |  |
| Health practitioners |  | End user |  |

**1e Scientific Abstract**

## Provide a summary of your proposal (topic, goal, approach and potential impact of results) in no

## more than 300 words.

Parkinson’s disease (PD) is a chronic neurodegenerative disease. The loss of dopamine producing neurons in PD patients results in motor and non-motor symptoms and there is still no cure. Currently, the diagnosis of PD is done by a holistic evaluation of the symptoms of a patient. This means the patient not can be diagnosed before the symptoms are visible and there is already a loss of dopamine-producing neurons before PD has been diagnosed. Also, misdiagnosis of PD is incredibly high. The goal of this study is therefore to come up with an innovative approach to diagnose PD with more accuracy and sooner. For this study this research group can use an existing data set. This data set contains anonymous data of health persons and persons diagnosed with PD. From these persons the timing of the key strokes of typing on a computer during their usual computer usage is recorded and stored anonymously and unreadable. A supervised machine learning approach might be possible to build a model to detect early symptoms of PD. A possible way to detect early symptoms of PD might be to differentiate between the kinematics of the keys presses on the left side versus keys pressed on the right side of the keyboard, since onset of the symptoms seem be unilateral. When this research groups achieves this goal, this model can be used in practices to diagnose PD sooner than with the current method. This might lead to PD patients getting medication sooner and thereby reducing and delaying the symptoms of PD.

### 1f Keywords

Parkinson’s Disease, diagnosis, fine motor skills, typing, machine learning

### 1g Scientific Background & Relevance to the research areas:

## Describe how the research fits to (at least three of the four) the research areas, use up to 700 words in total.

## Describe for each area the scientific research goal and specify for each area its relative scientific importance in the proposal and its relative contribution to the project. Which research area is scientifically most challenging in the proposal?

Parkinson’s disease (PD) is a chronic neurodegenerative disease. The loss of dopamine producing neurons in PD patients results in motor and non-motor symptoms and there is still no cure (Adams, 2017). Currently, the diagnosis of PD is done by a holistic evaluation of the symptoms of a patient (Sveinbjornsdottir, 2016). This means the patient not can be diagnosed before the symptoms are visible and there is already a loss of dopamine-producing neurons before PD has been diagnosed (Fearnley and Lees, 1991). The goal of this study is therefore to come up with an innovative approach to diagnose PD with more accuracy and sooner.

Symptoms in the motor control of PD patients such as bradykinesia, unilateral onset of the symptoms and persistency of these asymmetric symptoms (Sveinbjornsdottir, 2016), might have an effect on a simple typing task. Therefore, it might be possible to use key stroke data of typing on a computer to find differences between healthy and PD patients with mild symptoms.

For this study this research group can use the data of Adams (2017). This data set contains anonymous data of healthy persons and persons diagnosed with PD. This data set includes birth year, gender, presence of several symptoms, medicine usage, and PD severity. From these persons key stroke data on their own computer during their usual computer usage was recorded (using a custom key stroke recording program called ‘Tappy’) without any supervision. From the key strokes the date, timestamp, key on left or right side of keyboard, hold time, latency time, fly time and direction of the key pressing to the next key (left side to right side, left to left side etc.) was recorded.

Misdiagnosis of PD is incredibly high according to Singh and Xu (2019). Methods of data science might help in this case. A supervised machine learning approach might be able to build a model to detect early symptoms of PD. These symptoms might not be visible yet, however differences for instances in key stroke latencies, hold times and flight times might show differences between healthy persons and patients with mild PD.

A possible way to detect early symptoms might be to differentiate between the kinematics of the keys presses on the left side versus keys pressed on the right side of the keyboard, since onset of the symptoms seem be unilateral (Sveinbjornsdottir, 2016). When left hand versus right hand kinematics are not separately analysed, minor differences in for example key stroke latencies, hold times, and flight times between healthy subjects and patients with mild PD might not be detected.

The biggest challenge of the study is to find a model to that can differentiate between healthy and PD patients with mild symptoms with high accuracy and precision. When this research groups achieves this goal, this model can be used in practices to diagnose PD sooner than with the current method. This might lead to PD patients getting medication sooner and thereby reducing and delaying the symptoms of PD.

Data Science:

Sport Science:

Life Sciences and Health:

Creative Industry Research:

Social sciences and/or humanities (ethics and legislation):

**References**

Adams, W. R. (2017). High-accuracy detection of early parkinson's disease using multiple characteristics of finger movement while typing.*PloS One, 12*(11), e0188226.

Fearnley, J. M., & Lees, A. J. (1991). Ageing and parkinson's disease: Substantia nigra regional selectivity.*Brain : A Journal of Neurology, 114 ( Pt 5)*(5), 2283-2301.

Singh, S., & Xu, W. (2019). Robust detection of parkinson's disease using harvested smartphone voice data: A telemedicine approach.*Telemedicine Journal and E-Health : The Official Journal of the American Telemedicine Association.*

Sveinbjornsdottir, S. (2016). The clinical symptoms of parkinson's disease.*Journal of Neurochemistry, 139 Suppl 1*, 318-324.