

# Research plan



## Table of contents

<b>1</b>	<b>Introduction .....</b>	<b>3</b>
<b>2</b>	<b>Issue/research questions .....</b>	<b>Fout! Bladwijzer niet gedefinieerd.</b>
2.1	Trouble domain .....	4
2.2	Demarcation .....	<b>Fout! Bladwijzer niet gedefinieerd.</b>
2.3	Goal .....	<b>Fout! Bladwijzer niet gedefinieerd.</b>
2.4	Research questions .....	4
<b>3</b>	<b>Final product .....</b>	<b>Fout! Bladwijzer niet gedefinieerd.</b>
<b>4</b>	<b>Theoretical framework .....</b>	<b>6</b>
4.1	Start theoretical framework .....	6
4.2	Middle theoretical framework .....	6
4.3	End of theoretical framework .....	7
<b>5</b>	<b>Research design: Method of data collection .....</b>	<b>8</b>
5.1	Research design .....	8
5.2	Research conditions .....	8
5.3	Risk analysis .....	8
<b>6</b>	<b>Work planning .....</b>	<b>9</b>
	<b>Source list .....</b>	<b>11</b>

## 1 Introduction

Within this document you will find a global overview of our research plan. It contains the problem domain of our project that explains what problem the CBS is currently having and why our problem owner wants to have this subject researched. The global research plan also explains our current main research questions that is divided into multiple sub questions. By answering all sub questions, you can answer the main research questions.

Within the project we will be following the scrum methodology. You can find more about that in the 'planning' section of this document. If anything is not clear, or you want a better explanation please reach out to us.

## 2 Issue/research questions

### 2.1 Problem domain

#### First Iteration (by Matthew)

The Central Bureau voor de statistiek (CBS) has the wish to measure the intensity of physical activities. Currently they are using health surveys which are not **accurate** enough for their purposes. Therefore CBS has been looking into alternatives like the ActivPal accelerometer in combination with machine learning to give better and more accurate results when measuring the intensity of certain activities.

Because of this the CBS started to collect **lab** tests and started to measure the movements of 41 correspondents in their regular work week by using the Activepal Accelerator. It's our job to analyse, structure and build machine learning algorithms based on the collected data to see if we can determine if people adhere to (inter)national norm for physical activities and if we could measure the intensity of movement (without the heart rate information).

#### Second Iteration (by Adnan)

Statistics Netherlands (CBS) has the wish to see if their respondents are moving for at least 150 minutes per week in moderately intense physical activity.

Currently, they are measuring by asking their respondent or health surveys. The issue with this is that people are not very good at estimating the time they spent on moving and sport. This of course causes that they don't have very reliable data to work with. Therefore CBS has been looking into alternatives like the ActivPal accelerometer in combination with machine learning to give better and more accurate results when measuring the intensity of certain activities.

Because of this the CBS started to collect lab tests and started to measure the movements of 41 correspondents in their regular workweek by using the Activepal Accelerator. It's our job to analyse, structure and build machine learning algorithms based on the collected data to see if we can determine if people adhere to (inter)national norm for physical activities and if we could measure the intensity of movement (without the heart rate information).

### 2.2 Research questions

#### First Iteration (by Matthew)

Within the ActivePal project we will be working with two main research questions. To give a good answer to main research question. We have decided to split the questions up in different topics. The main questions are marked as bold, the sub questions are marked as Italic and the sub sub questions are marked as Underline.

1. **Can we use this knowledge to determine if people did their 150 minutes of moderate activity in the week data?**
  - a. *Can you recognize the lab activities in the week data?*
    - i. How do you recognize an activity in the lab data?
    - ii. What features do you need to learn a model to recognize activity's?
  - b. *Can you see patterns in the rest of the data that indicate doing sports?*
2. **How do we measure the intensity of movement in the ActivPAL data? (Without the heart rate information)**

**Commented [AA1]:** So, accuracy is the only problem of the CBS???

**Commented [AA2]:** And is the end goal of CBS to measure all activities in the lab only???

**Commented [AA3]:** ????

**Commented [AA4]:** It's not the point if you can recognize this, but if a computer algorithm can.

**Commented [AA5]:** 2.1 suggest you only have data collected during lab tests!!!!

**Commented [AA6]:** Is this enough?? Don't you want to know what sport or how active the person was?

**Commented [AA7]:** You're not measuring the movement in the data itself....

How can data science/machine learning aid in predicting the intensity of movement of a person wearing an ActivePal device in a lab situation, and is this estimation, based on ActivePal measurements only, accurate enough for the CBS to ...

I know you

- a. *Can you measure the intensity of movement by using a MET value of an activity?*
  - i. How do you calculate the MET value of an activity?
- b. **Can** *you measure the intensity of movement by measuring the Magnitude of acceleration of an activity?*
  - i. How do you calculate the Magnitude of acceleration value of an activity?

Every research question will be put into the Jira project management board, and the Jira Roadmap. With the corresponding tasks linked to it, by answering all questions and doing the corresponding tasks we can answer the main research questions.

## Second Iteration (by Ali)

Within the ActivePal project we will be working with three main research questions. To give a good answer to main research question. We have decided to split the questions up in different topics. The main questions are marked as bold, the sub questions are marked as italic.

1. **How can Machine Learning be used to predict the intensity of activities performed in a lab situation by a person, who is being monitored with Vyntus One and wearing ActivPal accelerometer?**
  - a. *What measurement does ActivPal use for intensity and why?*
  - b. *Is it possible to extract this intensity measurement values from just Vyntus One data, if so, how?*
  - c. *Is it possible to extract this intensity measurement values from just ActivPal accelerometer data, if so, how?*
2. **How can Machine Learning be used to predict the intensity of activities performed by a person wearing only the ActivPal accelerometer, based on the data gathered from Vyntus One and ActivPal accelerometer in the lab situation?**
  - a. *Are there one or more explanatory variables in the ActivPal accelerometer and respondents data for the intensity measurement values of main question 1, if so, what are they?*
3. **How can Machine Learning be used to determine whether people did their 150 minutes of moderate activity in ActivPal accelerometer data of an entire week?**
  - a. *How can Machine Learning be used to recognize the activities, performed in the lab situation, in the ActivPal accelerometer data?*
  - b. *How can Machine Learning be used to recognize patterns in the data that indicate doing sports?*

Every research question will be put into the Jira project management board, and the Jira Roadmap. With the corresponding tasks linked to it, by answering all questions and doing the corresponding tasks we can answer the main research questions.

## Third Iteration (by Matthew)

Within the ActivePal project we will be working with three main research questions. To give a good answer to main research question. We have decided to split the questions up in different topics. The main questions are marked as bold, the sub questions are marked as italic.

**Commented [AA8]:** So questions 2a and 2b are the only available alternatives, and you're going to ultimately choose one of them???

**Commented [T(9)]:** What machine learning can best be used to predict...

**Commented [T(10)]:** What would be the best model to do..

**Commented [T(11)]:** Duidelijk verschil maken tussen prediction van activiteit, intensity en classificatie.

**Commented [T(12)]:** Bij feedback duidelijk vragen wat je precies wilt weten (kwa feedback)

**Commented [T(13)]:** What kind of models could be used to predict intensity

**Commented [T(14)]:** How can hyperparameters be tweaked to predict the best (feature)

**Commented [T(15)]:** Meer in relatie plaatsen met hoofdvraag 1 en 2

**Commented [T(16R15)]:** Verwijzen naar de antwoorden van bijvoorbeeld eerdere vragen. Zoals: Intensity can best be measured by model x and so what would be the best features to use.

**Commented [T(17)]:** What machine learning can best be used to predict...

**Commented [T(18)]:** What would be the best model to do..

**Commented [T(19)]:** Duidelijk verschil maken tussen prediction van activiteit, intensity en classificatie.

**Commented [T(20)]:** Bij feedback duidelijk vragen wat je precies wilt weten (kwa feedback)

**4. How can Machine Learning be used to predict the intensity of activities performed in a lab situation by a person, who is being monitored with Vyntus One and wearing ActivPal accelerometer?**

- d. What measurement does ActivPal use for intensity and why?*
- e. Is it possible to extract this intensity measurement values from just Vyntus One data, if so, how?*

**5. How can Machine Learning be used to predict the intensity of activities performed by a person wearing only the ActivPal accelerometer, based on the data gathered from Vyntus One and ActivPal accelerometer in the lab situation?**

- b. What machine learning model can best be used to measure the intensity for each activity?*

**6. How can Machine Learning be used to determine whether people did their 150 minutes of moderate activity in ActivPal accelerometer data of an entire week?**

- c. How can Machine Learning be used to recognize the activities, performed in the lab situation, in the ActivPal accelerometer data?*

Every research question will be put into the Jira project management board, and the Jira Roadmap. With the corresponding tasks linked to it, by answering all questions and doing the corresponding tasks we can answer the main research questions.

## Theoretical framework

**Woordenaantal:** Er geldt geen richtlijn voor het aantal woorden van dit theoretisch kader. We adviseren om niet meer dan 3 A4'tjes hieraan te besteden.

Geef aan dat je in dit hoofdstuk de belangrijkste begrippen, theorieën en modellen die betrekking hebben op jouw onderzoeksonderwerp beschrijft.

### 3.1 Start theoretical framework

Bespreek in een korte alinea de opzet van je [theoretisch kader](#). Geef daarna antwoord op de volgende vragen:

- Wat zijn de kernbegrippen die relevant zijn voor jouw onderzoeksonderwerp?
- Waarom zijn deze relevant voor jouw onderzoeksonderwerp?

### 3.2 Middle theoretical framework

Hier bespreek je de theorieën en modellen die je nodig hebt om jouw onderzoek te doen.

- Welke theorieën en modellen zijn relevant voor jouw onderzoek?
- Waarom zijn deze relevant voor jouw onderzoek?

Formatted: Dutch (Netherlands)

Formatted: Dutch (Netherlands)

Formatted: Dutch (Netherlands)

Formatted: Dutch (Netherlands)

Field Code Changed

Formatted: Dutch (Netherlands)

Formatted: Dutch (Netherlands)

### 3.3 End of theoretical framework

- Welke vragen beoog je met jouw theoretisch kader te kunnen beantwoorden?

Let op: Ons artikel over het theoretisch kader gaat over een theoretisch kader in een scriptie. We raden je aan om te lezen wat de verschillen tussen de twee theoretische kaders zijn in ons artikel over het [Plan van Aanpak](#).

**Formatted:** Dutch (Netherlands)

**Field Code Changed**

**Formatted:** Dutch (Netherlands)

**Formatted:** Dutch (Netherlands)

## 4 Research design: Method of data collection (Mark)

**Woordenaantal:** Er geldt geen richtlijn voor het aantal woorden van de onderzoeksopzet. We adviseren om niet meer dan 2 tot 3 A4'tjes hieraan te besteden.

Geef aan dat je in dit hoofdstuk de [onderzoeksopzet](#), de randvoorwaarden voor je onderzoek en de risicoanalyse presenteert.

### 4.1 Research design

Bespreek het soort onderzoek dat je gaat doen.

- Doe je [kwalitatief](#) of [kwantitatief](#) onderzoek, of allebei?

Ga in op de wijze van dataverzameling.

- Ga je data verzamelen met behulp van [deskresearch](#) of [fieldresearch](#), of allebei?
  - Tot welke documenten heb je toegang nodig van welke tijdsperiode?
  - Hoeveel personen benader je voor de dataverzameling?
  - Wil je [enquêtes](#) verspreiden of [interviews](#) houden?
  - Etc.

Beschrijf welke data je gaat analyseren.

- Welke data ga je analyseren? Bijvoorbeeld bepaalde literatuur, enquêteresultaten of [transcripties](#).

Geef aan hoe je de data gaat analyseren, ofwel met welke methode(n).

- Hoe ga je de data analyseren?
  - Maak je bijvoorbeeld gebruik van [SPSS](#)?
  - Ga je de data vergelijken?
  - Gebruik je een bepaald model om de data te analyseren?
  - Etc.

### 4.2 Research conditions

Omschrijf de voorwaarden om de opdracht of het onderzoek uit te voeren.

- Wat heb je nodig om jouw opdracht of onderzoek uit te voeren?
  - Geld
  - Informatie
  - Tijd
  - Materiaal
  - Toestemming
  - Etc.

### 4.3 Risk analysis

Ga in op de risico's van de opdracht of het onderzoek en de maatregelen om deze risico's te vermijden.

- Welke risico's zijn er die je opdracht of onderzoek negatief kunnen beïnvloeden?
- Hoe kun je deze risico's voorkomen?

Formatted: Dutch (Netherlands)

Field Code Changed

Formatted: Dutch (Netherlands)

Formatted: Dutch (Netherlands)

Formatted: Dutch (Netherlands)

Formatted: Dutch (Netherlands)

Formatted: Dutch (Netherlands)

Field Code Changed

Field Code Changed

Formatted: Dutch (Netherlands)

Formatted: Dutch (Netherlands)

Field Code Changed

Formatted: Dutch (Netherlands)

Formatted: Dutch (Netherlands)

Formatted: Dutch (Netherlands)

Formatted: Dutch (Netherlands)

Field Code Changed

Formatted: Dutch (Netherlands)

Formatted: Dutch (Netherlands)

Field Code Changed

Formatted: Dutch (Netherlands)

Field Code Changed

Formatted: Dutch (Netherlands)

Formatted: Dutch (Netherlands)

Field Code Changed

Formatted: Dutch (Netherlands)

Formatted: Dutch (Netherlands)

Formatted: Dutch (Netherlands)

Formatted: Dutch (Netherlands)



## 5 Work planning

### First Iteration (by Matthew)

For the ActivePal project management we have been using the Agile (scrum) method. To follow the scrum method, there will be daily standups. Within our project the daily standup time is 9:30 and since day 1, we have been following this method every day. Within the daily standup we discuss the progress of the prior day and the plans for today. Also, we discuss issues that you might encountered. If you are stuck or not sure what you should be doing next, we make sure that clear that up and make new working pairs.

At the end of the sprint, we hold a retrospective. This is the best way to look back at the last two weeks and make an evaluation of the progress that have been made. Also, we use this time to see how everybody is feeling and if we need to make changes to make sure that everybody is happy.

Since we are using the Scrum method, this also means that every project member has a certain role. Every sprint we rotate the roles within our project group, to make sure that everybody is getting more experienced with Scrum and working together.

The infrastructure/tool we have been using for this is Jira. Jira is a plan, track and manage software that is mostly used for software development. Within the ActivePal project Jira is not only used for software development but also for the research development from week 1.

The main research questions have been divided into smaller tasks called research sub tasks. Every sub research question has a few tasks that help to answer the question and guide us to the right approach to give an answer. By combining Jira with our sub questions, we can put all the sub questions into the Jira Road Map. This gives the whole group a good overview of the questions that still need to be answered and the tasks needed to answer the sub questions.

To get access to our Jira board you can contact one of the ActivePal group members for an invite and a brief explanation.

### Second Iteration (by Colin, added to Matthew's part)

#### Procedure

For the ActivPal project management we have been using the Agile (scrum) method. To follow the scrum method, there will be daily stand-ups. Within our project the daily stand-up time is 9:30 and since day 1, we have been following this method every day. Within the daily stand-up we discuss the progress of the prior day and the plans for today.

#### Solving issues

Besides this, we discuss the issues that you've might encountered during your previous working day. If you are stuck or not sure what you should be doing next, we make sure to clear that up and make new working pairs. Usually when this happens there will be Teams calls after the stand-up to discuss or solve this problem with each other. This way the stand-ups won't take longer than they should be.

#### Sprint planning

Before holding our retrospective, we make sure to create a planning for the next sprint. During every sprint planning we, as a team, look at the current sprint on Jira. We discuss if we managed to complete the set sprint goal, if all the tasks are done, and what still must be done with the leftover tasks to complete them. The leftover tasks will be added to the new sprint.

**Commented [AA21]:** This chapter is now pure procedural. I miss the first setup of your product backlog. Chapter 5 should already have resulted in a lot of input for this!

**Commented [T(22):** Aantonen sprintplanning / aantonen sprints / Jira / retrospectives

Once the current sprint has been completed, we start the new sprint by specifying a goal for the new sprint. We usually look at the roadmap and what we have done the previous sprint to create a new goal. After agreeing as a team with the set goal, we have healthy discussions where everyone gets the chance to add new tickets to our backlog that will help achieve this goal.

#### Creation of the first sprint backlog

The creation of our first sprint backlog was a bit hectic. The project was new for all of us and nobody knew where to start or what the exact planning was going to be. We still managed to create a backlog for the first sprint which was mainly focussed on research.

Annemieke from CBS provided us with lots of useful papers to research. A task during the first sprint was filtering out useful papers, this way we did not waste time by rereading unusable papers. As a team we all picked different subjects which we thought would be beneficial for our project and created 1 ticket with small subtasks in it. Besides the different papers we also received different CSV files which contained the data we had to work with during the project. We created a ticket to research these CSV files and document as much as possible.

#### Retrospectives

After finishing a 2-week sprint we hold a retrospective to reflect and evaluate the process we've made over the past weeks. In these retrospectives we discuss what went well, what didn't go well, what each team member longed for and the actions we're going to take for the next sprint. During the retrospective we look back at the actions we wrote down from the previous sprint and see if we have worked on these actions. These retrospectives are a good tool to let every member on the team freely speak on how to feel about certain aspects over the last 2 weeks.

#### Project roles

At the start of the project, we agreed on switching roles every sprint. This would mean a new scrum master, communicator and note taker every sprint. This did not go as planned, we decided to keep the same scrum master for the entire project to make sure the project went smoothly. The communicator role was switched halfway through the project and note taker role was done by the whole team. Every member wrote down what was important for them and after every important meeting we discussed the previous meeting to make sure everyone is on the same level.

#### Planning tool

The infrastructure/tool we have been using for this is Jira. Jira is a plan, track and manage software that is mostly used for software development. Within the ActivPal project Jira is not only used for software development but also for the research development from week 1.

#### Research questions

The main research questions have been divided into smaller tasks called research sub tasks. Every sub research question has a few tasks that help to answer the question and guide us to the right approach to give an answer. By combining Jira with our sub questions, we can put all the sub questions into the Jira Road Map. This gives the whole group a good overview of the questions that still need to be answered and the tasks needed to answer the sub questions.

To get access to our Jira board you can contact one of the ActivPal group members for an invite and a brief explanation.

## Source list

Presenteer de bronnen in alfabetische volgorde met behulp van de [APA Generator](#) of [MLA Generator](#).

- Welke bronnen heb je gebruikt om dit Plan van Aanpak te schrijven?

**Formatted:** Dutch (Netherlands)

**Formatted:** Dutch (Netherlands)

**Field Code Changed**

**Formatted:** Dutch (Netherlands)

**Field Code Changed**

**Formatted:** Dutch (Netherlands)

**Formatted:** Dutch (Netherlands)