**Measuring Engineering Report**

**Task**

To deliver a report that considers the ways in which the software engineering process can be measured and assessed in terms of measurable data, an overview of the computational platforms available to perform this work, the algorithmic approaches available, and the ethics concerns surrounding this kind of analytics.

**Introduction**

The industry has a great interest in measuring and assessing software engineering processes. However this is not an easy task since the complexity of a software engineering task cannot be easily measured. There have been many models in the history of computer science in order to address this problem but the most reliable approaches are still depending on a great level of human assessment. The improvements in the field of artificial intelligence show a potential solution to this problem. However the information needed to operate these systems do come at a high computing and privacy cost. This report is taking a look into the questions what data can be used, what platforms can be used, which algorithms are available and the ethics concerns behind this kind of analytics.

**What data?**

Measuring the amount of work a software engineer puts into the development of a process is quite a complex task. One of the most basic approaches to this problem is to count up the number of lines type by the programmer in order to solve a problem. This approach is highly inaccurate for obvious reasons the number of lines used to solve a problem do not represent the problems level of complexity or the quality of the code produced by the programmer. A better approach is to measure the work done by a programmer in time. This gives a better overview of the complexity of the problem however it does not give any indication to the quality of the code and it also does not give any comparison to other programmers. Today many companies masseur their programming tasks in a score. This works in the way that every team member estimates the complexity of a given task in comparison with different tasks and gives them a score the average of the estimated scores than gets used as an indication of the tasks complexity. In that way the programmers get a feel for estimating there work and they can compare the work they do with their colleges within their team since the task has been scored by the entire team. This approach gives a much better estimate to the work done by an individual software engineer but it is very time consuming and skilled programmers need to spend a lot of time to estimate and maintain this process and it is all depending on the individual’s skill to collectively estimate tasks. So it is still of great interest for businesses to make the measuring and assessing of engineering processes more efficient and reliable. Many computer scientists believe that this can be achieved by collecting large amounts of date on every individual software engineer in order to draw up patterns in their behavior that indicate there productivity. So for example body movements like nodding or typing can be used to make assumptions to someone’s happiness or even their productivity by comparing them to a large data set. Of course this information does not give exact prediction´s to a individual however it might help to make faster decisions and also may help when looking at a large number of employs for example if large numbers of employees change their typing speed or their body language it might reflect on a change they experience in office. Collecting this range of personal date always comes with a security risk and most employees would feel unconfutable to work in an environment where your workstation or even wearable technology’s constantly collect data about them while they’re working. However the picture the people have about privacy is rapidly changing over the last few years and therefore it could just be a matter of time that employees in the future will accept this kind of transparency in their workplace. Team work is also an important aspect of productivity people that work in a group collectively tends to outperform individual s. A key aspect in team work is communication so chatrooms or GitHub commits can be a good indication how a team cooperates and therefor uses their full potential as a group. In the future this could be even extended to sensors that take in the level of communication within an office since real life human interactions are very important for the levels of happiness within an office. Especially the amount of conversation taking place during the breaks can be a good indication for the happiness of the employees.

**Where to compute?**

At the moment the amount of data held about an employee in a company is small enough to be handled locally in most firms. However this could rapidly change once the amount of data hold on an individual software engineer increases. In order to make complex conclusions it would be necessary to compare a worker to a large data set and it also involves a lot of computing power in order manage these data sets. Only very large firms could afford to maintain their own large server systems in order to process there employee’s information. This could result in the outsourcing of the data processing. Companies could than simply collect the data and hand them over to the datacenters that process them these data centers could than use there collected information’s to make up statistics and draw conclusions to their employees and how the perform within their team. This obviously comes with a security risk. Handing out personal information to another company is a privacy issue but with increasing data it might be the only way smaller businesses can compete with large companies.

**What algorithms?**

Over the years many people tried to make up a formula or an algorithm that would be able to calculate or measure productivity. However these algorithms no matter how complicated they are always have limitations and most time just give an estimation on the productivity of a person. They are limited by the amount of data gathered on the person but also there is no improvement once a formula is build an regular algorithm is not able to reflect on previous inputs in order to specify on there conclusion, but this is really important especially when we are interested in the productivity of individuals. Artificial intelligence can be a solution to this problem. Every time an AI gets new input it learns so it is constantly improving this has a high potential however AI’s need a lot of computing power in order to be affective and the training process of the neuronal network is especially difficult since it has to train on very large data sets which are simple not available yet. Training an neuron network to measure software engineering process would mean to feed the program relevant information and then giving the program a conclusion after time the neuron network will be able to predict the conclusions by strengthening or weakening the connections between the neurons, but these conclusions all have to be done by humans and also a large percentage of the data also has to be gathered by humans this means that the training process does not scale very well and a neuron network that needs to make complex conclusions needs a lot of training. So maybe it is not even possible for a human to train an AI very complicated processes simple because it would take too much time to train it and we first have to change the training process so that the network starts to training itself. Once the training problem of neuronal networks is solved AI is going to show its full potential this could mean that AI’s with enough computing power could solve highly complex tasks like the measurement of productivity of an engineer.

**Ethics**

The measurement and assessing of software engineers also includes an ethnic conflict. On the one hand the benefits of measuring the productivity of a software developer makes a company more productive and therefor more competitive however it comes at a cost. Employees could lose a large amount of privacy up to the point where we find our society in a “Big Brother” like situation. The computing power of modern data centers increases rapidly and we most likely will get to a point where we have the capabilities to process as much data of employee as we are able to gather. The question is will this improve our society and what impacts will it have. Of course everybody likes privacy and no one likes the feeling of being watched, but the way we view privacy has made already major changes over the last few years. The increase in social media has already shown an impact on our view what we see as private information we care about and what not. In fact there are already large amounts of data gathered on users in order to place advertisement effectively or for security reasons and most people do not seem to have a major problem with it. We know that people are willing to make sacrifices to the marketing industry for the benefits of social media so way should they not be willing to sacrifice privacy at the workplace in order to improve the economy. Once the workers will feel a benefit in giving up information it could quickly become a standard to which the people get used to. But this does not answer the question whether or not this improves our society we are already really productive and this technology could simple just be used to increase our advantage over poorer countries which cannot afford this kind of technology. We are also already living in a quite materialistic society where everybody has more than they need. Therefor increasing the productivity in producing non-essential products may not be worth the sacrifice of privacy. There are also other ethical issues with this topic like reliability the more trust we put in these technology’s the more we are depending on them to work without faults. It is easy to see how the miss judgment of a machine can lead to bad consequences when we trust the machine to judge on the importance of a person in a workplace and many people would not feel comfortable by knowing that such important topics are handled by a computer. However we could also argue that a machine brings benefits to judging people a machine does not discriminating because of race or gender and it also would be more consistent in its decisions than a human. No matter how we argue in favor or agent measuring employees by using personal date the capabilities of machines are constantly improving and once we have new capabilities due to technology we tend to use it. So it might be more in our interest to try influencing its developing process rather than fighting the idea completely.

**Philip Albrecht**

**14316854**