Product vision

Group Health informatics-2

Rick Proost – 4173619
Pascal Remeijsen – 4286243
Wim Spaargaren – 4178068
Arnout vd Knaap – 4223969
Jelmer de Boer - 4223152

Contents

Introduction

Researchers want to analyze raw data, but this is not always structured in a way it can easily be entered into static analysis tools. You get a diversity of data types like text and excel files from real world applications that needs to be formatted.

This product will format and do quick analytics in an intuitive way, so researchers will be able to do their research in an efficient manner, and later use tools like *SPSS(www-01.ibm.com, 2015)* to do further analysis over multiple users. SPSS is a static analysis tool made by IBM, which allows a user to use certain analysis techniques on a formatted data-set. The data that our product is meant for, is gathered from self monitoring patients who just got a new kidney using the method described in *Zelfmonitoring van nierfunctie na niertransplantatie: de patiënt als regisseur*(2014). This project is called the ADMIRE Project *(http://ii.tudelft.nl/admire/, 2013)*. In short, a study where patients are given tools to monitor their situation following a kidney transplant, and they have to collect data on a sensor, and on a website. The device used for gathering the data at home was tested in this study *Application of a point of care creatinine device for trend monitoring in kidney transplant patients: fit for purpose?(2015)* Another study was conducted for how the visualisation of the system has to be in : *Extracting usability information from user interface events(2015).*

It is important that good tools are developed for the processing of this kind of data because it will save a lot of time for the researchers and potentially doctors as explained in *(Schneeweiss,2014)*. During the development we used the book *Interaction Design(2007)* as a reference to ensure we deliver a product which is pleasant to interact with.

Product

This project will be made with the vision of a flexible data processor in mind. We want our product to take raw data, and give the user a way to work with this data efficiently.

This product will be made with the primary client being researchers from the ADMIRE project, but the product will be as flexible as possible. The product could be used by other researchers, who need a way to extract certain aspects of a defined data-set. This product will not be used during the research, but it will be used after the completion to parse the gathered data. This is not a tool to do statistical analysis. It will rather be a bridge between the raw data and the static analysis tool. The product could be used to format the data so it can be used as input for a static analysis tool, or to get visualizations for simple questions about patients. Thus, the target audience is people who need to process structured data before they use static analysis tools on it, or only have simple questions that could be visualized by our product.

The product will address a number of customer needs, which are not always possible with static analysis. Namely, the pre-processing of the data. This will include 8 basic transformations of sequential data analysis : chunking, comments, codes, connections, comparisons, constraints, conversions and computation as explained in *(Sanderson & Fisher, 1994)*. Not all static analysis tools are able to do this, so we will take care of these operations. To enable the user to effectively apply transformations to the data, we will create a language which will allow for easily defining the order and specifications of the

transformations on the data. For optimal use, the product will also enable the user to define input and output file locations, as well as define the input and output format of the data. This last feature will make the product useful to all kinds of users, besides the researchers of the ADMIRE project. Besides these operations to the data, we will also enable the user to take the processed data and make certain simple visualizations showing relations between-, or the values of sets of data. The goal is to offer a way of intuitively exploring the gathered data.

The language created for efficient use of the tool will be at the heart of the product. It will allow users to customize their queries and define the format of the input as well as the output. This satisfies the customer need for flexible data transformations. The visualization features will satisfy the need for simple data exploration. If the client has substantial statistical analysis requirements, they should use our tool to transform the data, but do the analysis in a separate tool. But if the user needs are simple queries, the product is able to show a variety of visualizations for the data.

The project will be realized in the fourth quarter of the university-year. This will span roughly eight weeks. There is no monetary budget, but the five members of the project-group working on the product, will invest considerable amounts of time and effort to ensure a satisfactory product gets delivered.

Interface

To be able to better convey our vision we include a sketch of what our user-interface might look like:
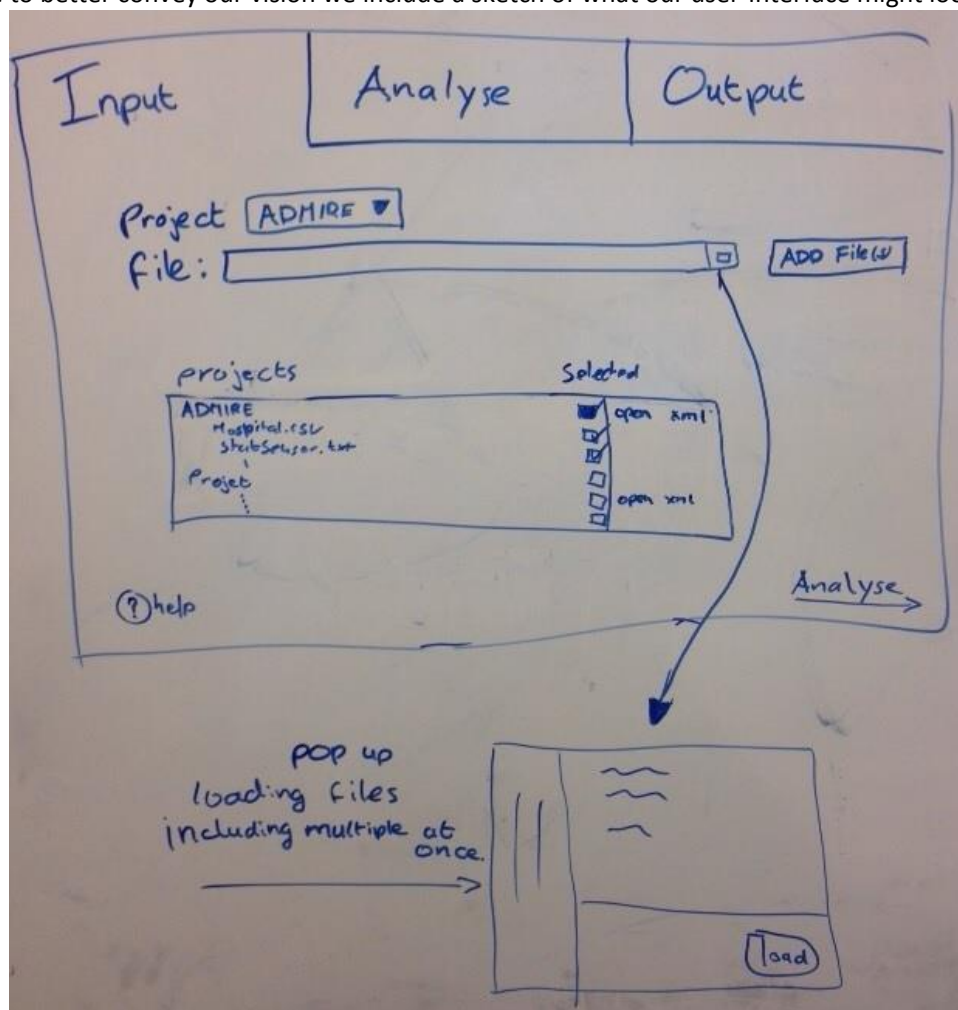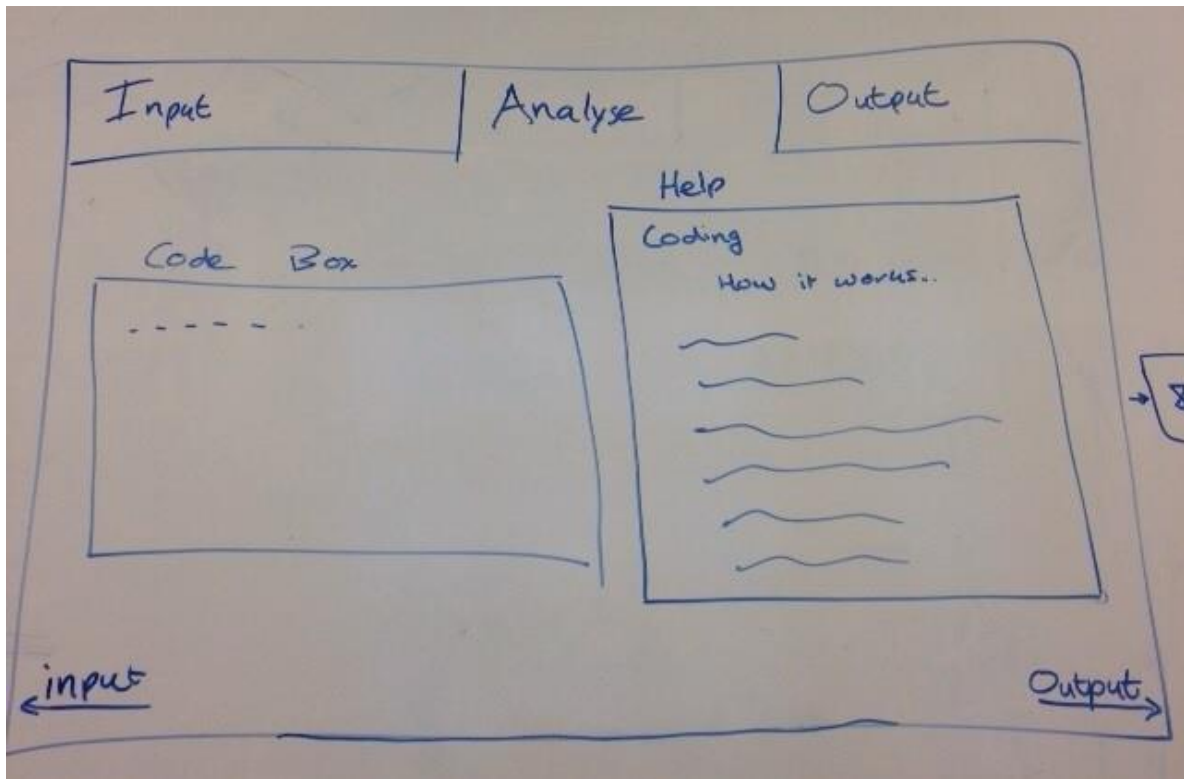

Figure 1: InputTab sketch
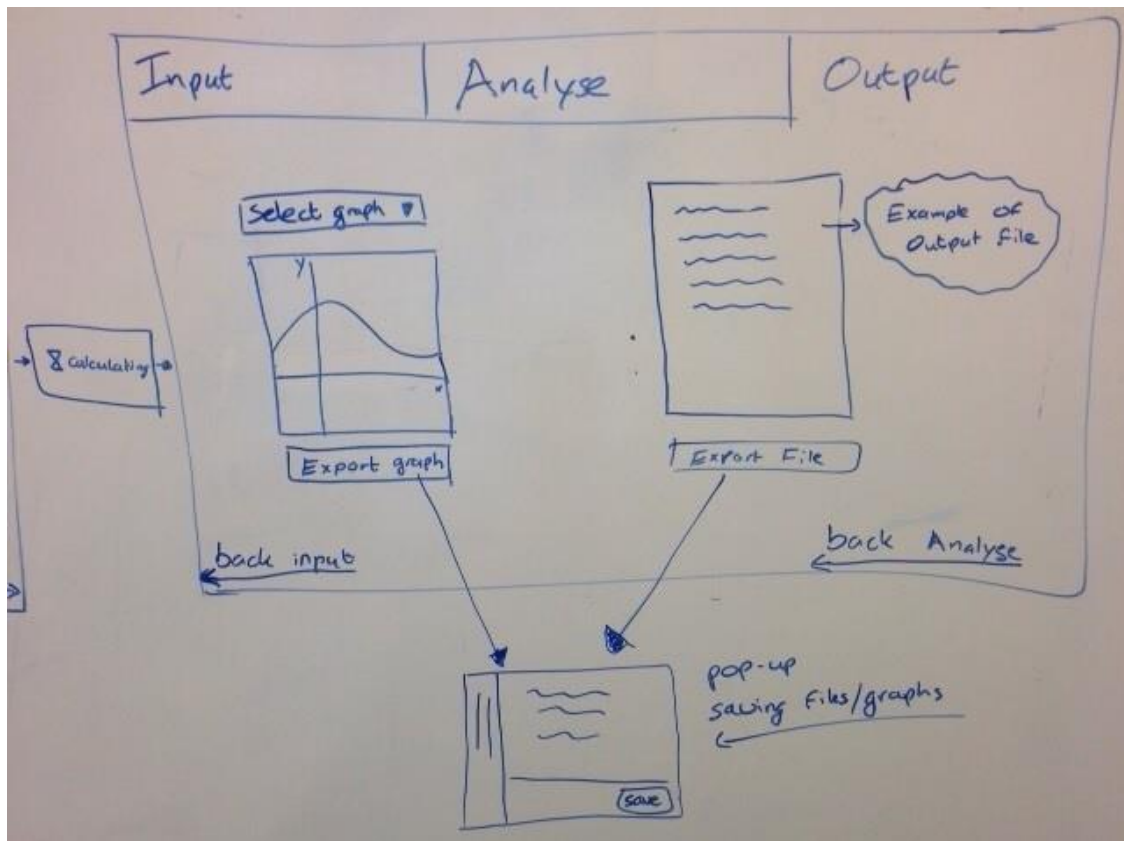
Figure 2: AnalyseTab sketch

Figure 3: OutputTab sketch

Product vision

An elevator pitch could be : **FOR** Researcher **WHO NEED** a way of transforming sets of raw data into a format which can be used as input for statistical analysis tools. **ADDITIONLALLY** we would like to make the target audience bigger by allowing the input- and output- formats to be defined by the user. This way it can be used by other users who work with data sets besides researchers. The product IS A tool which can apply certain transformations to a data set **WITH THE GOALS OF** formatting data, or extracting some meaning from it. **UNLIKE** statistical analysis tools, we cannot apply statistical operations to the data, we just transform it into a format which is easier to use in further exploration of the data. **OUR** Services also include making certain data visualization of the processed data set possible.

MoSCoW

In this section we display the features of the program, described according to the MoSCoW method.
This means that every feature/aspect of the program are divided into four groups:

Must haves: Features whose functioning is required to get the program working. These features have a high priority.

Should haves: Features that are wished for, but in absence they do not harm the program. They have a medium priority.

Could haves: Features that are only implemented if there is time, but otherwise are planned for a follow-up project. They have a low priority.

Won't haves: Features that will not be implemented in this project.

Must haves

- Program is able to format a .txt output file for analysis tool SPSS(sprint 3).
- Program is able to read multiple files which could be compared to the data of the ADMIRE project(sprint 3)
- Program is able to filter data (constraints). (sprint 3)
- Program is able to compare values of input files. (sprint 4)
- Program is able to chunk on data sets (sprint 4)
- Program is able to run an algorithm or input standard values to be compared with actual data. (sprint 5)
- Program is able to encode certain values (codes).(sprint 6)
- Program is able to count how many times values/events occur (computation) (sprint 6)
- Program is able to add comments to the data set output (comments). (sprint 7)

Should haves

- Program can output a user defined data set (sprint 6)
- Program is able to connect different values as one event (connections). (sprint 7)
- Program is able to create new values from actual data set (conversions). (sprint 8)
- Program is able to output if patients measure more if other factors like blood pressure increase (sprint 7)

Could haves
- Program is able to export visualizations of the imported data (sprint 7)
- Program can output if external factors (holidays) affect how patients measure (sprint 8)

Won't haves
- Program can do static analysis over more than one person.

Reference list

- Sanderson, P. M., & Fisher, C. (1994). Exploratory sequential data analysis: Foundations. *Human–Computer Interaction*, *9*(3-4), 251-317.
- Céline van Lint, Paul van der Boog, Paul Schenk, Sandra van Dijk, Fred Romijn, Christa Cobbaert. *Zelfmonitoring van nierfunctie na niertransplantatie: de patiënt als regisseur.* Voorjaarscongres Nederlandse Vereniging voor Klinische Chemie en Laboratoriumgeneeskunde (NVKC) 9-11 april 2014, Veldhoven.
- Hilbert, D. M., & Redmiles, D. F. (2000). Extracting usability information from user interface events. *ACM Computing Surveys (CSUR)*, *32*(4), 384-421.
- ADMIRE. (2013, September 11). Home page of the ADMIRE project. Retrieved from *http://ii.tudelft.nl/admire/*
- van Lint, C. L., van der Boog, P. J., Romijn, F. P., Schenk, P. W., van Dijk, S., Rövekamp, T. J., ... & Cobbaert, C. M. (2015). Application of a point of care creatinine device for trend monitoring in kidney transplant patients: fit for purpose?. *Clinical Chemistry and Laboratory Medicine (CCLM)*.
- Schneeweiss, S. (2014). Learning from big health care data. *New England Journal of Medicine*, *370*(23), 2161-2163.
- SPSS-software (IBM SPSS software) *http://www-01.ibm.com/software/nl/analytics/spss/*
- Sharp, H., Rogers, Y., and Preece, J. (2007). Interaction design: beyond human-computer interaction (2nd ed). Chichester, England: John Wiley & Sons (ISBN 978-0-470-01866-8)