Product vision

Group Health informatics-2

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

Preface

This document describes the vision of the product for Context Product: Health informatics 2.
The vision of the product defines how the product will be structured and implemented. The main features and final product is envisioned here but will still be subjected to changes.

Table of contents

Introduction

Researchers want to analyze raw data, but this is not always structured in a way it can easily be put in 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 easy manner.

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. 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 query's 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.

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.