Visual Analytics: Dataset Proposal for Assignment

Saumya Roy[*]   Dido Verstegen[†]   Siddharth Sekar[‡]

1 INTRODUCTION

Visual Analytics is the science of analytical reasoning supported by a highly interactive visual interface. Our objective for the assignment is to utilise the 'Disappearance of the Gastech' dataset. By using visualisations, we can quickly identify patterns, trends, and relationships within the data making it easier for us understand, explore, and communicate insights. The Dataset provides information about the disappearance of employees of Gastech, a Tethys-based company, that primarily works on natural gas production in the island of Kronos. The disappearance takes place amidst a company celebration in January of 2014. Other information regarding this disappearance is to be investigated from the data that is provided. The dataset primarily consists of employee records, a map of Kronos, employee emails and other such documentation. It also consists information on the protectors of Kronos, an environmental activist group that is primarily suspected to be behind the disappearance.

1.1 Data and Task Analysis

With the dataset provided, we need to understand the visualisations to understand the type of information we can obtain from them. Primarily we intend to formulate the following questions that we think will help us in understaing this mysterious disappearance.

- Which members of Gastech went missing and why did they go missing?

- Who are the Protectors of Kronos (POK) and how are they involved in the disappearance?

- Is there any link between Gastech and the POK?

- What is the primary mode of communication between the employees and what can we analyse from the messages sent through it?

- What information can we gather from the new reporting that happened in Kronos prior to and during the disappearance?

- What can we say about the individual employees of Gastech?

- Is there any specific area of importance in the map of Kronos?

We believe these questions will provide us with useful insight for creating our visual encodings. It would also help us in co-relating and increasing the interactivity of the the visualisation tool.

1.2 Visualisation Tool Platform

We intend to use Python based Dash framework and Plotly library to implement the visual encodings mostly because of their primary features such as interactivity and flexibility. It is also easy for Integration and Collaboration since it is Python based and can be easily hosted and shared via URLs. Dash and Plotly also have an active

[*]s.roy1@student.tue.nl

[†]d.c.r.verstegen@student.tue.nl

[‡]s.sekar@student.tue.nl

community base that ensures that users can get help with any issues they encounter.
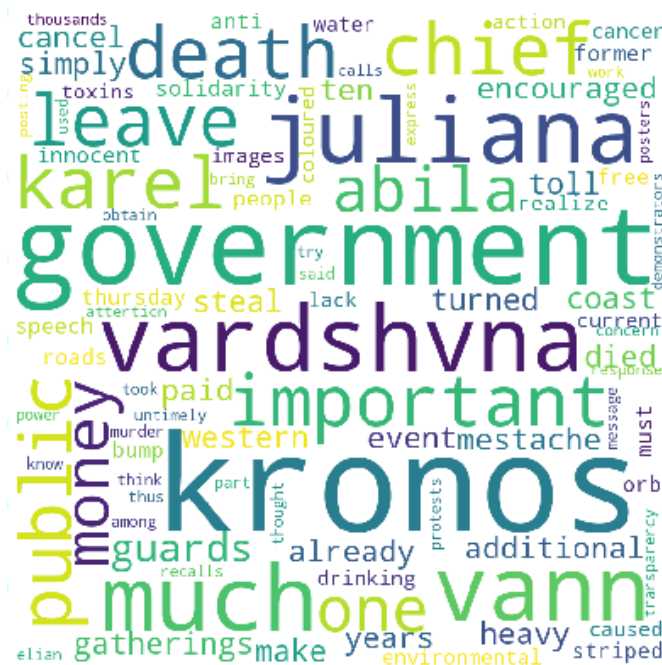
Fig1: Potential words of interest from the articles