Data Systems Project

...

Ana Chaloska, Anisha Sloeserwij, Danny de Vries, Faye Tervoort, Jesse Kommandeur
University of Amsterdam
Amsterdam, The Netherlands

ABSTRACT
KEYWORDS

drugs, google trends, court cases, Netherlands, visualization

1 INTRODUCTION

For the police, it is important to monitor the modus operandi of criminal organizations. By analyzing the modus operandi, the police can gain insight into the plans of the criminal organizations in order to intercept them. Indeed, the modus operandi describes the entire way criminal organizations operate, from the beginning of the production stage to the use of its consumer.

This report examines a particular stage of the modus operandi in general in The Netherlands and presents the results in a visual way so that the stakeholder, the Police Academy, can get an insight on this stage of the modus operandi. Various trends in the field of drugs in the Netherlands were researched. Using a google search trends analysis and an analysis to court cases in the Netherlands, an attempt was made to make connections between these two sources. These results are presented in a visual way that provides new insights for the Police Academy.

1.1 Problem statement

2 RELATED WORK

2.1 Stakeholder research

A study of the stakeholder with their values and what they would like to see as an outcome from us.

The stakeholder is the Police Academy and the number one goal they want from us is to gain insight into the 'Modus operandi' of criminal organizations in the field of drugs. The values and norms of the Police Academy must be clear and the result must be designed accordingly.

We can represent this well with a visual that states what the stakeholder's values and norms are, what they want to see as a result and what their focus points are.

2.2 Modus operandi

Telling about modus operandi, what it entails and perhaps some research that is already done on 'modus operandi'. Telling about what stage of modus operandi we want to investigate and how that can contribute for the new insights of the Police Academy.

2.3 Drugs

What are drugs? Which drugs do we want to focus on? Why?

2.4 Drugs in the Netherlands

When did some drugs become legal or illegal, when did a certain law come through that, for example, had a major impact on the use of (a certain) drug, etc. Perhaps depict this in a timeline which shows the events of influence (and then describe it in the text).

2.5 Requirements

What are the requirements that we need in order to make the product. Make a list where all requirements are mentioned, like

2.6

2.7

3 METHODS

3.1 Explain why we use these methods and why they are useful for our case, mention what Ana Barros told us

3.2 Google trend research

3.3 Court cases from rechtspraak.nl

3.4 News articles

3.5 Requirements

These user requirements are derived from the case description provided by the client as well as feedback from the stakeholder during the initial ideation and prototyping phase of the project and where further narrowed down during the project design workshops using the MoSCoW priorization method. The term 'user' in the requirements refers to two specific types of similar target audiences that will make use of our prototype, police agents who want to explore the dataset to gain insights and data analysts who want to filter and compare our datasets.

(1) (M) The user must be able to use the prototype on a personal computer and interface with a screen
(2) (M) The user must be able to filter the datasets to compare different years of data
(3) (M) The user must be able to filter the datasets to compare different types of drugs
(4) (M) The user must be able to overlay multiple datasets and trend lines on top of eachother
(5) (M) The system uses open-source software and not locked-in corporate data tools
(6) (M) The system has a user-friendly visual design and interaction design
(7) (S) The user should create an account to store specific and personalized filters
(8) (S) The user should be able to download the raw datasets in specific file formats

(9) (S) The user should be able to navigate between different overviews showing corresponding data

(10) (C) The user could upload their own dataset and sources of specfic drugs and news sources

(11) (C) The system uses real-time up-to-date API data

4 RESULTS

4.0.1 Prototype (dashboard).

Based on these requirements and the datasets a custom web-based visualization was created to allow the analyst to visualize explore, compare and filter the datasets. When filtering the charts update in real-time with keyframed animations. Screenshots and a live version of the prototype included in appendix 2a. For this prototype demo the user is logged-in with a default user account which is shown in the sidebar navigation. From there the user has four different overview pages to navigate to.

Events overview page

The first is an events overview page which shows a line chart with on the x-axis the week numbers per year and a relative score from 0 to 100 on the y-axis for our three datasets, the Google Trends, NOS News data and Twitter tweets. On the left bottom is a legend with a list of drug-related events. The user can click on an event which will highlight the week in which the event occured in the line chart to more clearly visualize peaks in the data which indicate event-based drug popularity. A small description of the event, most popular drug and event dates also is shown. The bottom-right shows the Google Trends data in a relative score over all 5 years of our dataset. This shows more clearly shows the lack of peaks of events in covid pandemic years.

Region overview page

The second page is a chloropleth map page showing the regions of the Netherlands. Each region is color coded using a linear color scale based (blue interpolation) on the relative score of Google searches for that region. On this page the user also has the option to focus on years and drugs but also on specific event dates. This for example shows that in Flevoland when the Lowlands event is happening there are alot of searches in that region for the drugs XTC.

Related queries page

The third page is the related queries page which shows a line chart for our three chosen drugs as filter options. The y-axis is a relative score from 0 to 100 and the x-axis are the week numbers for our specific year. The user can filter between different years. On the bottom are three polar charts which show the related drugs people are searching for with a relative score and a subset of the drugs in the legend. For example, people who search for cocaine also search for crack very often. It also shows that before 2019 GHB was not searched for by many people. It increased in popularity after that year.

Settings page

A settings page is also included. This page shows the different datasets used in the dashboard and a download button which allows the user to download a specific dataset either as .csv or .json to store on their computer.

4.0.2 Web application.

The prototype is a web-based application using web standards and open-source software and libraries. It being web-based allow the dashboard to run operating system independent. Only a webbrowser installed on the users device is required. The application will mainly be used in a desktop environment by the analyst so the dashboard is not fully responsive and not mobile optimized. The web application is created with the open-source front-end framework Svelte [1] and UI framework SvelteKit which allows the application to be build in components, each chart is rendered seperately making it more efficient to add functionality (add datasets, render different chart types) in the future but also makes the dashboard performant when more data and charts are added since Svelte already pre-renders the page offloading work from the browser. Working in components and with a frameworks such as Svelte allows future web developers to get up and running fast and add more functionality in a progressively enhanced manner. Svelte can be downloaded as a module (package) from NPM [2] and uses the JavaScript back-end run-time Node.js [3].

For the charts the JavaScript charting library Chart.js [4] is integrated into the components which allows charts to be rendered in HTML5 Canvas without much configuration. With Chart.js you can add a specific dataset and the scales of the axis will automatically change accordingly to the scale defined by the data. The filter options and updating of the charts is more custom, it uses JavaScript utility functions to allow the data of the to be preprocessed and have only the data changed not the scales of the whole charts. For the map page an additional Chart.js Geo Plugin is used to render the Chloropleth map. It uses a TopoJSON file to render the regions of the Netherlands and filters the properties within that .json file to render a score for each of the region.

The source code for the dashboard is published open-source on GitHub using the MIT license. A live version of the dashboard is contionously deployed on hosting platform Netlify [5]. Corresponding links can be found in appendix 2b.

5 CONCLUSION

5.1 Give an answer to the research question

6 FUTURE WORK

6.0.1 Expand data sources.
A subset of drugs and years was used for our research. To allow for more exploration and filterability the dashboard could be expanded with more datasets. Especially more relevant drugs (upcoming drugs, non-legal substances, NSPs) could be added as filter options. We also currently only use NOS news data but other more popular dutch news sources (e.g. RTL nieuws, Nu.nl) could be scraped to give a more accurate representation of the popularity of drugs mentioned in news articles. This will both improve the quality of the dataset by more accurately calculating a relative score as well as quantitify by gathering multiple sources and aggregating.

6.0.2 Real-time API data.
Currently the dasboard relies on exported data that is then loaded into the web visualization. As a

---

[1] https://svelte.dev
[2] https://www.npmjs.com
[3] https://nodejs.org/
[4] https://www.chartjs.org
[5] https://www.netlify.com

futher enhancement for the prototype and to make it more dynamic is to have the dataset exposed as an API which the dashboard can then fetch getting up-to-date real-time data. In this beta version datasets need to be added manually to the GitHub repository.

6.0.3 Usability testing. Further usability testing needs to be done to validate the User Experience (UX) and User Interface (UI) of the prototype to more accurately represent the needs of the police analyst. Basic user testing was done on a small group of people and feedback from the client was incorporated. But still, a lot of assumptions about the user have been made. User testing the live version of the prototype to a larger user base will uncover hidden problems. The feedback from the users would further validate the workings of the prototype.

In general further research is required to conclude that online data-gathering tools fare good indicators for predicting the popularity of specfic drugs.

7 DISCUSSION
8 APPENDIX