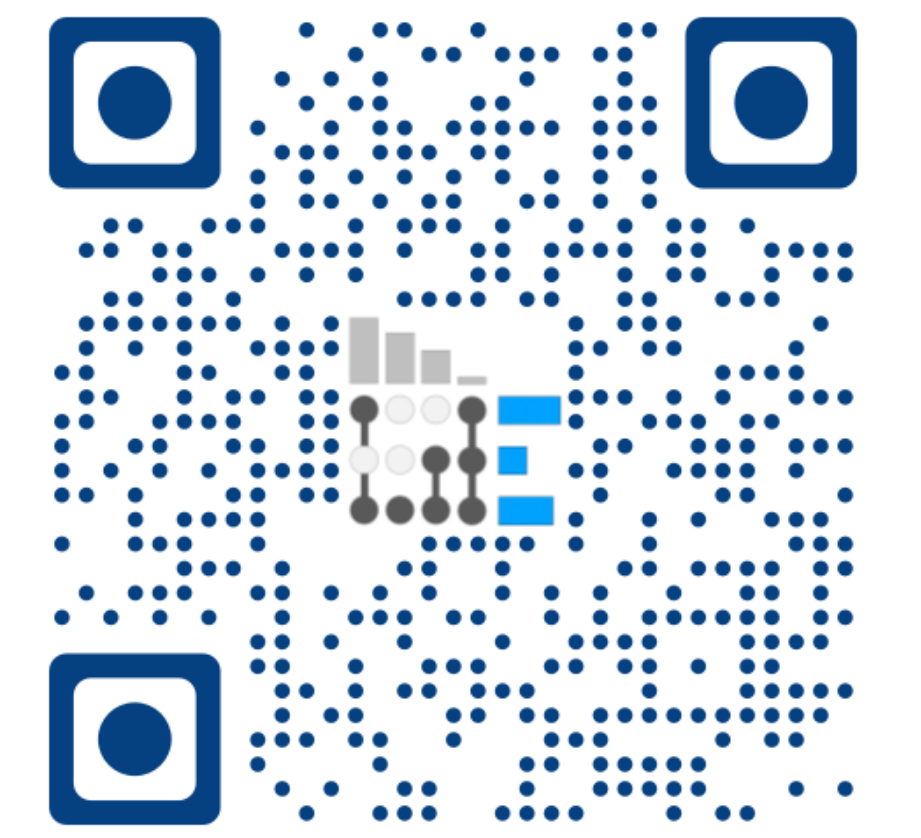


Enhancing Accessibility of UpSet Plots with Text Descriptions

Ishrat Jahan Eliza, Jake Wagoner, Jack Wilburn, Nate Lanza, and Daniel Hajas

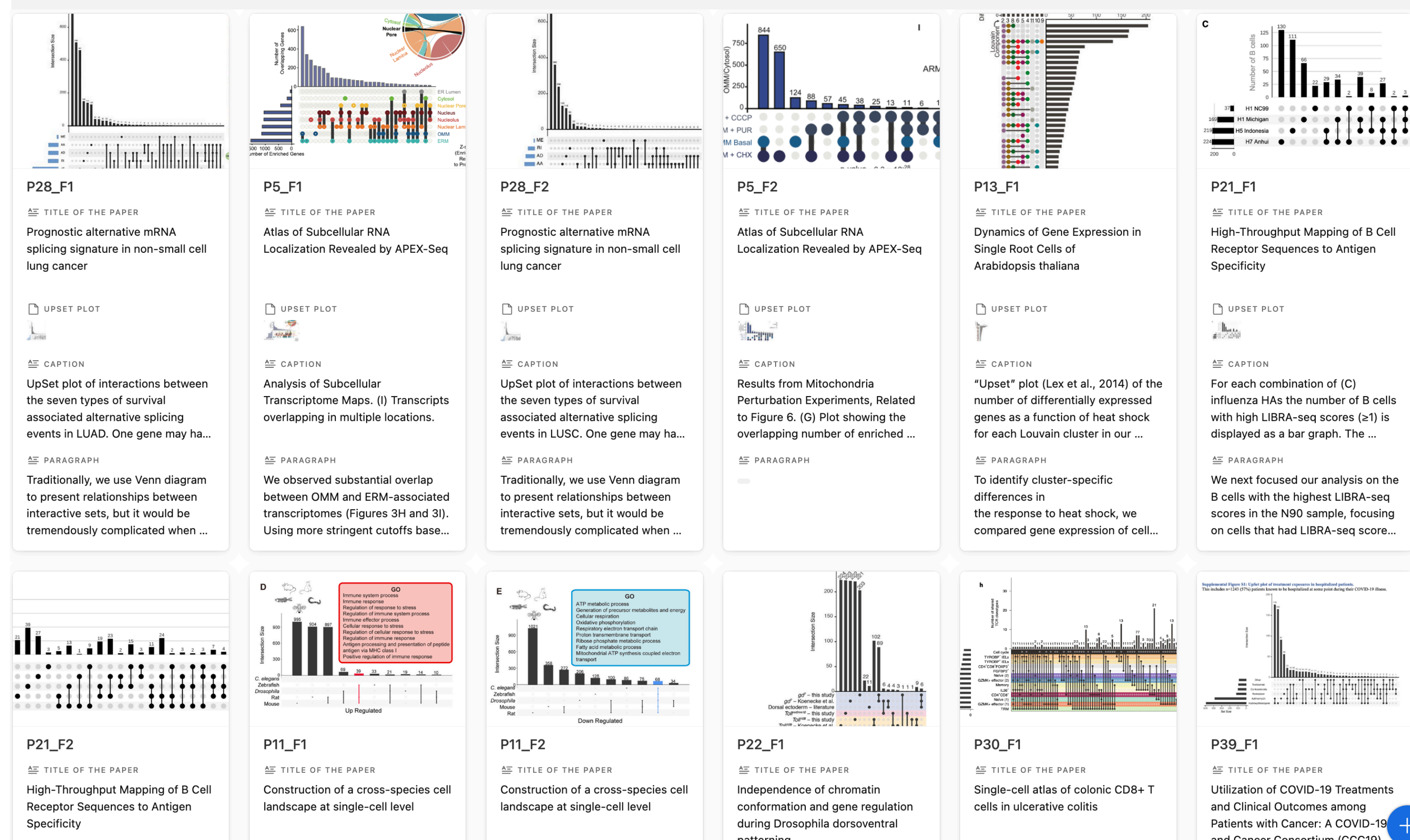


Data visualizations are typically not accessible to blind and low-vision users. The most widely used remedy for making data visualizations accessible are text descriptions. Generating text descriptions for complex visualizations is challenging because the patterns observed in these plots have not been sufficiently described.

We use UpSet plots as a representative complex visualization and generate text descriptions for it.

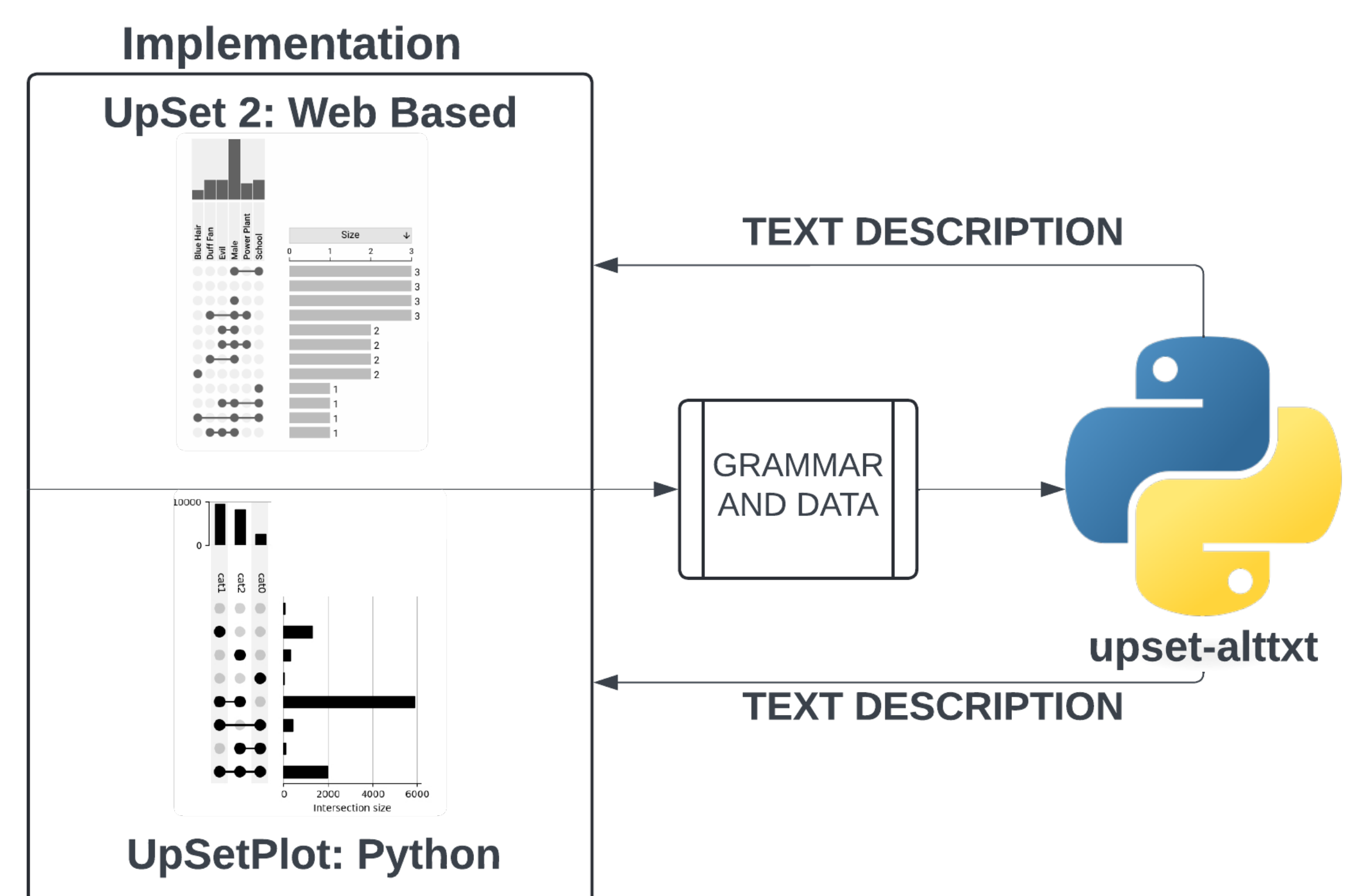
Identifying Patterns in UpSet Plots

79 plots from 40 published papers



Text Description Generation

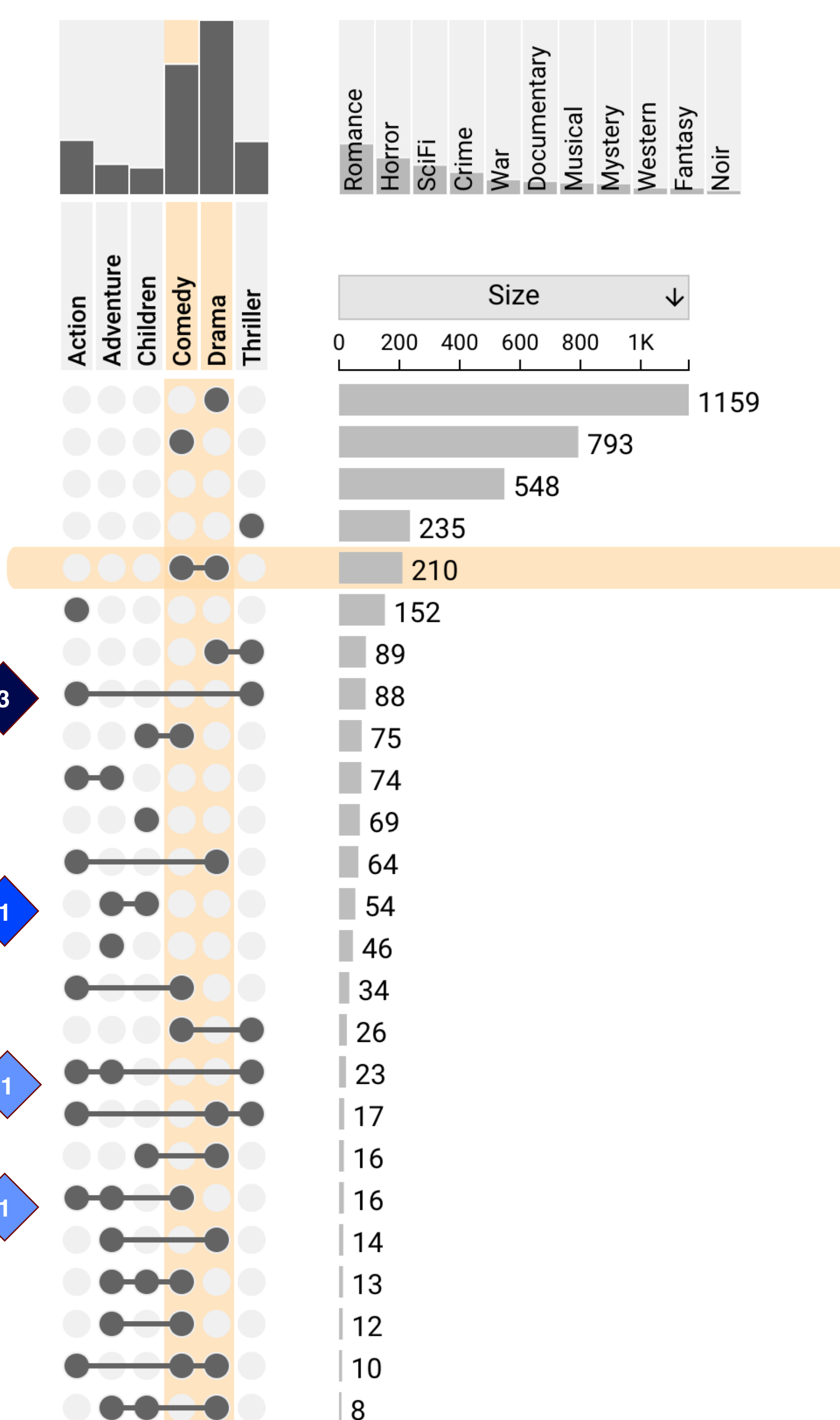
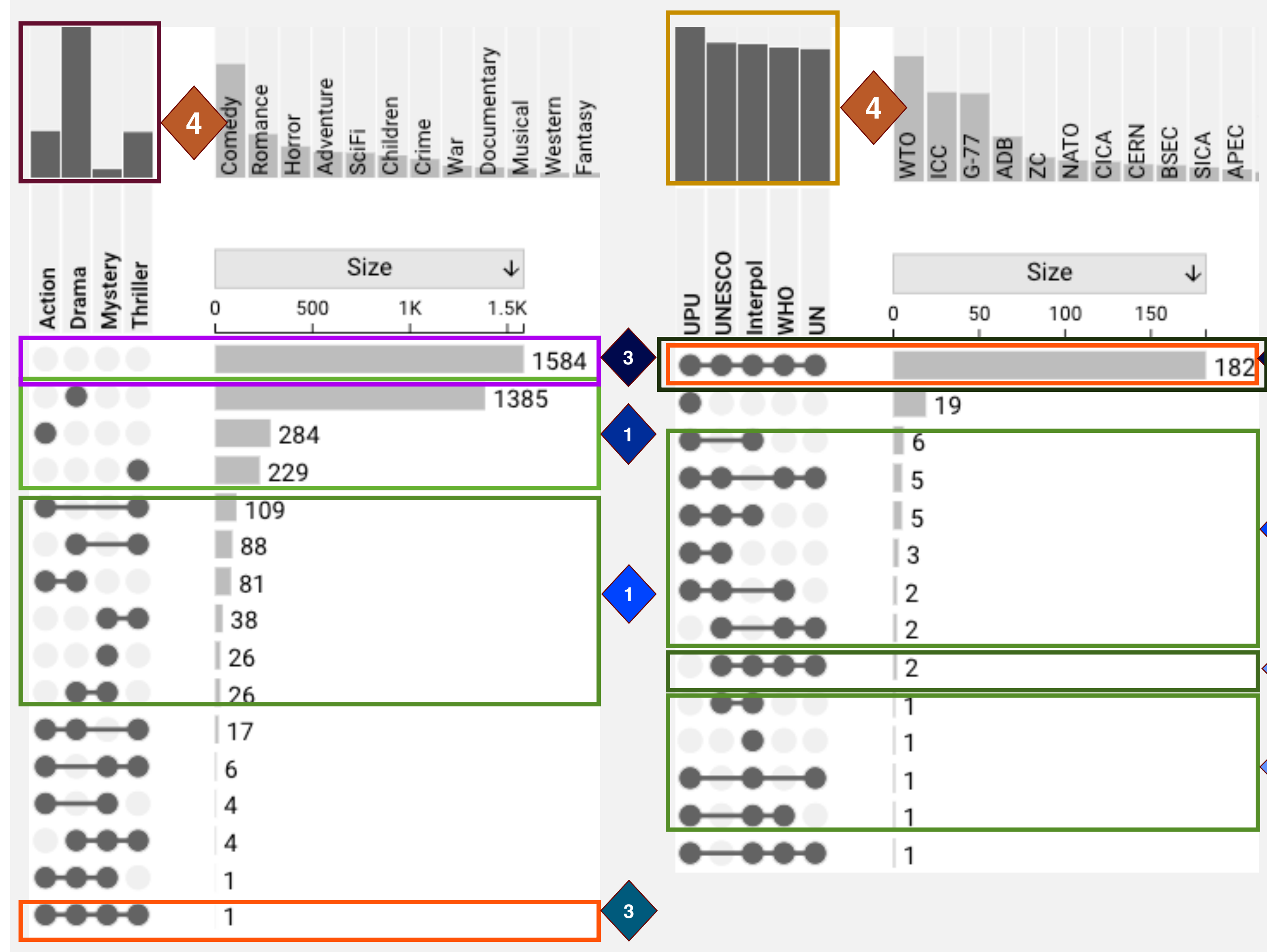
Our custom text generation package takes a configuration and a dataset and returns text descriptions: (1) a 2-3 sentence short description, and (2) a long description.



- 1 **Intersection Patterns:** Are there many **independent, low-degree, medium-degree** and/or **high-order** set intersections?
- 2 **Intersection Size Patterns:** Are the intersection sizes **small/medium/large/largest** associated with the intersection patterns?
- 3 **Specific Attributes:** Is the **all-set intersection** and/or **the empty intersection present/largest**?
- 4 **Set Sizes:** Are the set-sizes **roughly-equal/diverge a bit/diverge a lot**?

Using Text Descriptions

We present our text description in UpSet 2, an interactive web-based UpSet implementation. Additionally, we are able to generate text descriptions in UpSetPlot, a popular Python implementation of UpSet.



Dataset Properties

The dataset shows attributes of movies and their genres. The dataset contains 17 sets and 6303 elements, of which 6 sets are shown in the plot.

Set Properties

The set sizes are diverging a lot, ranging from 251 to 1603. The largest set is Drama with 1603 movies, followed by Comedy with 1200, Action with 503, Thriller with 492, Adventure with 283, and Children with 251.

Intersection Properties

The plot is sorted by size in descending order. There are 37 non-empty intersections, all of which are shown in the plot. The largest 5 intersections are Just Drama (1159), Just Comedy (793), the empty intersection (548), Just Thriller (235), and Drama, and Comedy (210).

Statistical Information

The average intersection size is 104, and the median is 16. The 90th percentile is 235, and the 10th percentile is 1. The largest set, Drama, is present in 40.5% of all non-empty intersections. The smallest set, Children, is present in 29.7% of all non-empty intersections.

Trend Analysis

The intersection sizes peak at a value of 1159 and then drastically flatten down to 1. The empty intersection is present with a size of 548. An all set intersection is not present. The individual set intersections are in largest and large intersections. The low degree set intersections lie in small and medium sized intersections. The medium degree set intersections can be seen among small and medium sized intersections. No high order intersections are present.

Acknowledgement: We gratefully acknowledge Kiran Gadhve, and funding by the Chan Zuckerberg Initiative through the EOSS5 program.

learn more about UpSet plot: <https://upset.app/>



visualization design lab