# Calliope·Data: An Intelligent Visual Data Story Generation Platform for Mass

Nan Cao*  Danqing Shi†  Yanqiu Wu  Xinyue Xu  Fuling Sun  Yi Guo  Nan Chen
Mengdi Sun  Xingyu Lan  Pei Liu  Siji Chen  Ligan Cai  Yang Shi  Qing Chen

Intelligent Big Data Visualization Lab
Tongji University

## ABSTRACT

Calliope·Data[1] is a novel visual data story generation platform that employs advanced AI techniques to automatically analyze data and represent data insights in the form of narrative visualization. It helps non-expert users create visual data stories from an input spreadsheet through an automatic process without any technique barrier. It also facilities the easy revision of the generated story based on an online story editor. To better support easy sharing and communication, Calliope·Data further incorporates a public website with social features such as story exploration, sharing, rating and comments. The evaluation shows the usability and usefulness of the platform.

**Index Terms:** Human-centered computing—Visualization—Visualization tools and systems

## 1 INTRODUCTION

A visual data story is a series of meaningfully connected story pieces, known as data facts, represented in the form of narrative visualization. To generate a visual data story, one needs to master three key techniques: data analysis, visualization, and scripting. To lower the technique barrier, our previous work proposed an intelligent system, Calliope [1], to generate visual data stories directly from an input spreadsheet in a fully automatic process.

In this poster, we extend the prototype design of Calliope to an online intelligent visual data story generation platform, Calliope·Data, which incorporates an efficient generation algorithm, an online story editor, and a public social website. In particular, the platform utilizes the generation algorithm to explores a large story design space given by the input spreadsheet and create a meaningful and informative data story [1]. The online editor provides comprehensive and flexible interactions to facilitate users to edit the generated story. To support easy communication, the public social website supports social features such as story exploration, sharing, rating, and comments. Furthermore, we also introduce Calliope·Lite[2], a lite version of Calliope without social features for quick story generation and privacy security.

## 2 CALLIOPE·DATA

The main activities on the Calliope·Data platform are to upload and manage data, generate and edit stories, and communicate the visualization results. Each part is described in the following subsections.

### 2.1 Data Management

To create a data story, users upload spreadsheets via an HTML form. Currently, Calliope·Data supports the spreadsheet in CSV format.
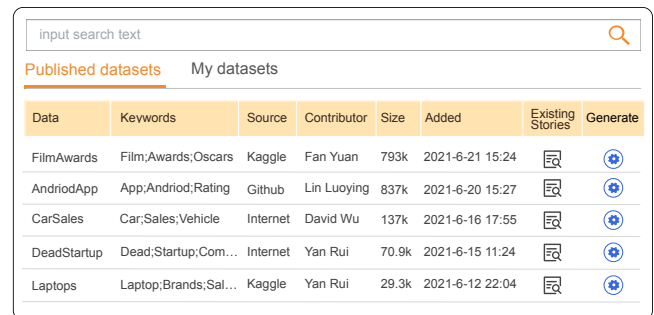
---

*e-mail:nan.cao@gmail.com
†e-mail:sdq@tongji.edu.cn

The upload form contains several text fields where the user can input the data title, keywords, and source. This information can help other users to understand and reuse these uploaded datasets.

The next step is data pre-processing for enabling further story generation. After a user has uploaded a spreadsheet, Calliope·Data will automatically label all data columns into four types, including *numerical*, *temporal*, *categorical*, and *geographical*. Then, Calliope·Data transforms each data column into a unified format according to its type. For example, the temporal column will be automatically transformed into the "year/month/day" format.



| Data | Keywords | Source | Contributor | Size | Added | Existing Stories | Generate |
|------|----------|--------|-------------|------|-------|------------------|----------|
| FilmAwards | Film;Awards;Oscars | Kaggle | Fan Yuan | 793k | 2021-6-21 15:24 | | |
| AndriodApp | App;Andriod;Rating | Github | Lin Luoying | 837k | 2021-6-20 15:27 | | |
| CarSales | Car;Sales;Vehicle | Internet | David Wu | 137k | 2021-6-16 17:55 | | |
| DeadStartup | Dead;Startup;Com… | Internet | Yan Rui | 70.9k | 2021-6-15 11:24 | | |
| Laptops | Laptop;Brands;Sal… | Kaggle | Yan Rui | 29.3k | 2021-6-12 22:04 | | |

Figure 1: The data browse page on Calliope·Data, giving an overview of all datasets currently uploaded to the site.

All the uploaded datasets are listed on the data browser page (Fig. 1). Users can browse the recently uploaded datasets or directly use the search box to find the specific dataset. Each item in the list provides the overview information of a dataset, including title, keywords, source, contributor, data size, and uploaded date. Each item also provides a link that shows all existing stories generated based on this dataset. User can further manage their own datasets or use the public dataset to generate a new story.

### 2.2 Visualization

A data story can be generated by clicking the "generation" button on the data page. The platform will then trigger the story generation engine to generate a sequence of data facts from the input spreadsheet. As the search space of a spreadsheet is exponential to the number of data columns and data attributes in each column, Calliope·Data uses the logic-oriented Monte Carlo tree search algorithm to explore the huge search space logically to guarantee the generation efficiency [1]. The time spent on each searching step can be configured by the user. Thus, the story generation can stop within a time limit to ensure efficiency. The generated data story will be visualized as a series of captioned visualization charts.

The generated story can be further edited in the story editor. As shown in Fig. 2, the story editor consists of three major views: the storyline view (Fig. 2-3), the configuration view (Fig. 2-2), and story visualization view (Fig. 2-1). In the storyline view, the generated data facts are shown in a row (Fig. 2-3), in which a user can remove a fact or change the generated narrative order based on his/her own preferences. Each fact is visualized by a chart and captioned by a
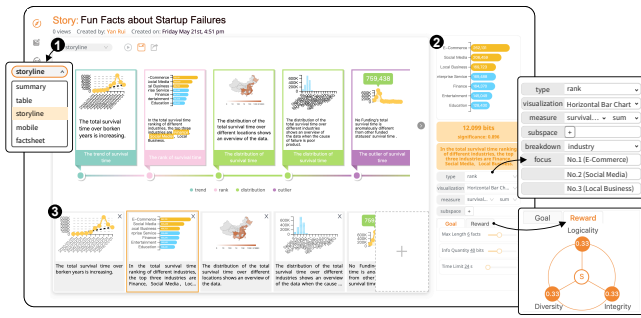
Figure 2: The story editor of Calliope·Data platform consists of three major views: (1) the storyline view for storyline editing, (2) the configuration view for fact editing and story generation, and (3) the story visualization view for the visual data story preview and sharing.

generated text description. When a fact is selected, the data details on each of its fact fields and visual and textual representations can be edited in the configuration view (Fig. 2-2). The user can also set the story generation goal and adjust the reward function in a group of configuration panels for the story regeneration. The generated data story can be visualized in the story visualization view (Fig. 2-1). In the end, a user can save the final result for further sharing.

## 2.3 Social Features

A data story gallery (Fig. 3) is integrated in the Calliope·Data to showcase the data stories created by various users based on different datasets. Each story item in the gallery includes the story title, preview, creator, rating, and sharing times. Users can click an item and enter the data story page to view the whole visual data story. For users who are interested in a specific topic, Calliope·Data provides a text-based search box for searching published data stories.
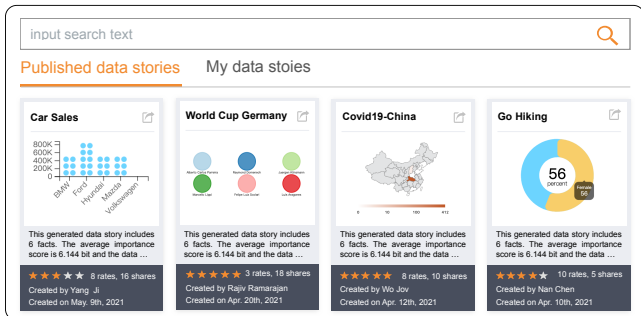


Figure 3: The gallery page contains a collection of visual data stories crerated by various users based on different datasets.

Users can share their own data stories or other public stories easily by clicking the "share" button. It will export a web page that other users can view the story via a public web link. In a specific data story page, users can feedback the story using the 5-star rating. The top-rated stories will be considered as popular stories, which will be displayed on the homepage of the platform. Furthermore, users can communicate with each other using textual comments. They can leave comments on the story or directly reply to another comment. This feature can engage the author and viewers in a discussion about the visual data story.

Shared data stories can be presented among three different modes in the story visualization view via a drop-down menu, including storyline, swiper, and factsheet for different application scenarios (Fig. 4). The storyline and swiper visualization modes are respectively designed to facilitate an efficient story exploration on tablets and smartphones. In particular, the captioned charts showing data facts in the story are horizontally aligned in a row to represent the

narrative order in the storyline mode and are shown one at a time in the swiper mode. A user can swipe the touch screen to explore the story through these representations. Finally, the factsheet mode is designed to show the story in the form of a poster that could be easily printed out.
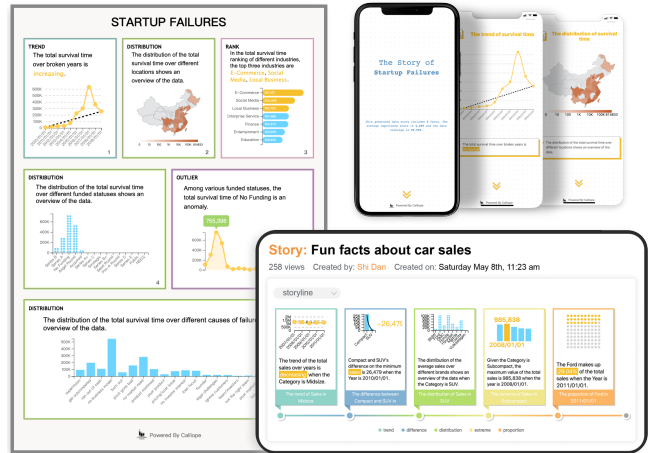


Figure 4: Calliope·Data provides three visualization modes to share the generated data story: (1) storyline, (2) swiper, and (3) factsheet.

## 2.4 Lite Version

To support the efficient story generation and privacy security, we also provide Calliope·Lite, a lite version of Calliope·Data without the on-site social features. As a user uploads the data file, Calliope·Lite will instantly analyze the data file and show the results in a chosen visual narrative form. Then, the user can further edit the result or directly download it as a PDF file to support off-site communication. After generation, Calliope·Lite automatically destroys all uploaded files from our servers.

## 3 EVALUATION AND FUTURE WORK

To evaluate the usability and usefulness of Calliope·Data in practice, we present an overview of activities on the platform. Since the launch of Calliope·Data, it has received 3750 views and collected 1893 user activities every month. One hundred twenty-eight registered users have uploaded 233 datasets and create 175 various visual data stories, which are shared 337 times in all. Users made 90 comments on the platform. About 45% of the comments give positive feedbacks on stories, and other comments lead to the discussion about the story. Several authors commented that with the help of the platform, they could quickly create a story by first automatic generation and then revise it accordingly. These results indicate that Calliope·Data can ensure the quality of generated stories and support easy communication on the platform.

To conclude, this poster presents Calliope·Data, a novel platform designed for automatically generating visual data stories from a spreadsheet. The platform incorporates a public social website to support story exploration, sharing, and communication. In the next step, we aim to involve more general users to use Calliope·Data and collect feedback to improve the efficiency and the user experience. Besides, we also plan to support more diverse data types, such as graphs, spatio-temporal data, or textual data. Overall, we believe Calliope·Data will power the future of data-driven storytelling.

## REFERENCES

[1] D. Shi, X. Xu, F. Sun, Y. Shi, and N. Cao. Calliope: Automatic visual data story generation from a spreadsheet. *IEEE Transactions on Visualization and Computer Graphics*, 2020.