Authoring Logically Sequenced Visual Data Stories with Gravity

Humphrey O. Obie[c,*], Caslon Chua[a], Iman Avazpour[b], Mohamed Abdelrazek[b], John Grundy[c], Tomasz Bednarz[d]

[a]*Swinburne University of Technology, Melbourne, Australia*
[b]*Deakin University, Geelong, Australia*
[c]*Monash University, Melbourne, Australia*
[d]*CSIRO's Data61, Sydney, Australia*

Abstract

Visual data stories have been shown to significantly aid the comprehension and short-term memorability of statistical facts and value messages. Hence, they are a promising medium for communicating complex information to a target audience, fulfilling the communicative goal of information visualisation. However, creating visual data stories often requires a plethora of tools to execute the visual data story creation process - creating visualisations, developing a logical connection between visualisations, and presenting the story based on the meaningfully sequenced visualisations. Current information visualisation tools either focus primarily on exploration or lack sequencing models for visualisation presentation, even though transitioning and sequencing have been shown to affect user understanding and interpretation of visualisations. We present Gravity, a system that consolidates the different phases of the visual data story creation process, and recommends and visualises logically sequenced sets of interactive visualisations to support the presentation of coherent narratives and visual data stories. We also report on an evaluation study with representative participants. Our results show that participants learned to use the system within a short time

---

[*]Corresponding author

Email addresses: humphrey.obie@monash.edu (Humphrey O. Obie), cchua@swin.edu.au (Caslon Chua), iman.avazpour@deakin.edu.au (Iman Avazpour), mohamed.abdelrazek@deakin.edu.au (Mohamed Abdelrazek), john.grundy@monash.edu (John Grundy), tomasz.bednarz@data61.csiro.au (Tomasz Bednarz)

and successfully created visual data stories with minimal guidance.

1. Introduction

Stories have been an essential medium for the communication of information in human societies [1]. They convey information in a way that is both interesting and engaging, and link events in a coherent whole, making them memorable [2]. Hence storytelling media have been well utilised in the fields of journalism, literature and film [3, 4, 5, 6]. This appropriacy of storytelling coupled with the current development of multimedia technology have created an upsurge in visualisation research interest in integrating stories and information visualisation - a subfield otherwise known as narrative visualisation or visual data storytelling [7, 8]. There has also been an increased interest in data journalism that has led to communication-oriented visualisation essay platforms such as Kiln's Flourish[1], The Pudding[2], and Distill[3].

Narrative visualisation (visual data stories), i.e., the combination of information visualisation with storytelling mechanisms, have been shown to significantly aid the comprehension and short-term memorability of statistical facts and value messages in a target audience [9]. Moreover, visual data stories grounded in facts have the potential to change people's beliefs and attitudes towards matters under consideration and, consequently affect their decision-making process [10]. Therefore, visual data stories are a promising media for communicating complex information to a target audience in a comprehensible and memorable way, thus fulfilling the communicative goal of information visualisation [8].

An important element of comprehensible visual data stories is the sequencing of the visualisation pieces that constitute the visual data story. An optimal sequence of visualisations for a linear visual data story is a sequence that conveys

---

[1]`https://flourish.studio/`
[2]`https://pudding.cool/`
[3]`https://distill.pub`

2

the data in a clear and logical manner [11], and consequently reduces the cognitive effort required to understand the visual data story, e.g., chronologically ordered visualisations, causal transitions between visualisations where the presentation of one visualisation follows the other to show a causal relationship [12].

Nevertheless, creating visual data stories is not an easy task. It demands substantial time and effort. It also often requires a variety of tools to execute the different phases of the visual data story creation process; creating visualisations, developing a logical connection between visualisations, and presenting the story based on the meaningfully sequenced visualisations [13]. Information visualisation tools (e.g., Voyager [14], Keshif [15]) largely cater to the exploratory aspect of the visual data story creation process but provide little support for the presentation goals of information visualisation. On the other hand, more recent tools (e.g., Ellipsis [1], Dataclips[16]) feature some presentation capability. However, they do not incorporate transitioning and sequencing models for presenting visualisations, considering that such sequences have been shown to affect user understanding and interpretation of visualisations [17, 12].

The aim of this work is the consolidation of the different phases of the visual data story creation process in a single tool that supports the exploratory and presentation goals of information visualisation, and integrate and visualise models for logically sequencing sets of interactive visualisations to support the development of coherent narratives. To this end, we present *Gravity*, a system that supports the creation of visual data stories. It adopts a meta-visualisation approach [18] in creating visualisation sequences either manually or using the system's sequence recommender. Gravity treats a visual narrative as a graph and encodes individual interactive visualisations as graph nodes. Directed edges between nodes represent logical connections between graph nodes, i.e., visualisations. In this visual narrative graph, pairs of nodes with the least sequence cost are connected to each other to form a visual narrative. Gravity supports the creation of interactive visualisations and visualisation sequences, and annotation of nodes with insight notes for reference and presentation purposes. In addition, the system also provides an overview of prior-created visualisations,

3

the inspection of individual visualisations, support for ethical forms and recording of presentation using a basic presentation slides deck. These features all support the goal of visual data story creation.

The design of Gravity is based on an abstract task analysis and consideration for the visual data story creation process. The system supports asynchronous collaboration and review of prior sessions via file creation, import and export. To evaluate Gravity, we conduct a user evaluation study with 10 representative users. The results show that participants successfully created visual data stories with minimal guidance and provide insight into how the system supports users through the visual data story creation process.

To summarise, our contributions are fourfold: (1) an abstract task analysis design for creating visual data stories; (2) consolidation of the different phases of the visual data story creation process in one system; (3) a meta-visualisation of interactive visualisation sequences; and (4) development and evaluation of the Gravity system.

The rest of this paper is organised as follows: we relate a motivating scenario in section 2. We discuss related literature in section 3. In section 4, we discuss our adopted design approach. Section 5 discusses the system and components design. In section 6, we discuss our user evaluation study results, and potential threats to validity and the control we applied. Finally, we conclude the paper and also outline direction for future work in section 7.

2. Motivating Scenario

In practice, visual data stories authors (such as data analysts) use a variety of tools to create visualisations, identify sequences of interesting story pieces, build their presentations, and deliver their data story to their target audience [16, 13]. This process is iterative and involves going back and forth between tools to create compelling and engaging narratives. However, this workflow involves a back-and-forth process between multiple tools, limits flexibility, and can be expensive in time and effort [19]. For instance, once a visualisation has

4

been exported from a visualisation authoring tool to a presentation tool, it loses connection to the data, since exporting is a one-way process [20], and the changes made in the presentation tool cannot be propagated back to the visualisation authoring tool because they are not dynamically bound together. In light of this, we describe below a motivating scenario of a typical data analyst whose job entails carrying out analysis and giving presentations based on the data and results of their analysis.

Consider a hypothetical scenario involving Matt, a data analyst. Matt's role involves working with datasets and crafting compelling narratives that are grounded in data. He is currently tasked with the responsibility of producing a visual data story video based on the *2019 economic freedom*[4] dataset, showing the different continents' economic freedom status based on the current index. Figure 1 captures Matt's visual data story creation process using 3 different tools to achieve his goal.

Matt accesses a visualisation authoring system and loads the *2019 economic freedom* dataset. He decides to first explore the relationship between *Trade Freedom and Tariff*. He selects appropriate visual encodings and creates an interactive scatterplot (Figure 1a). He then adds a suitable title to the visualisation and exports it as a static image file for later use in his visual data story presentation. He creates and exports 4 subsequent visualisations for presentation as well.

Matt loads the previously exported visualisations into a presentation software. He chooses to add insight notes to each visualisation for reference purpose while he is giving the presentation later. However, he also needs to confirm some underlying data points and details for the visualisations before adding the note, and because the current visualisations in the presentation software are static files, he has to go back to the source visualisation authoring system to complete

---

[4]Index of Economic Freedom measures economic freedom of 186 countries based on trade freedom, business freedom, investment freedom, and property rights. Data source: `https://www.heritage.org/index/`
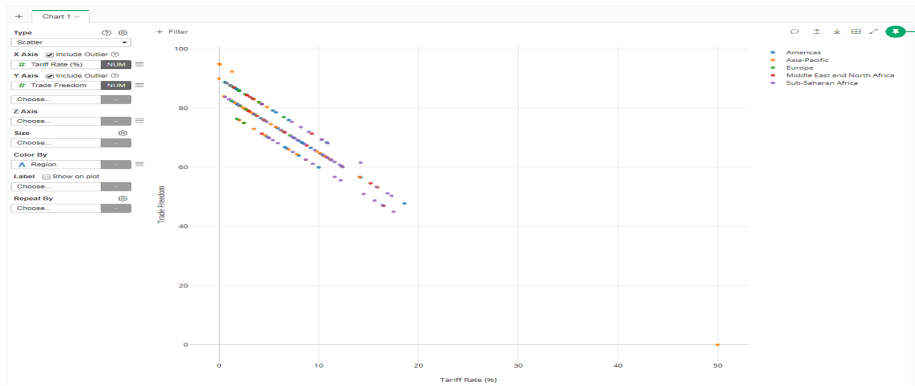
(a) Creating visualisation story pieces using a visualisation authoring tool.
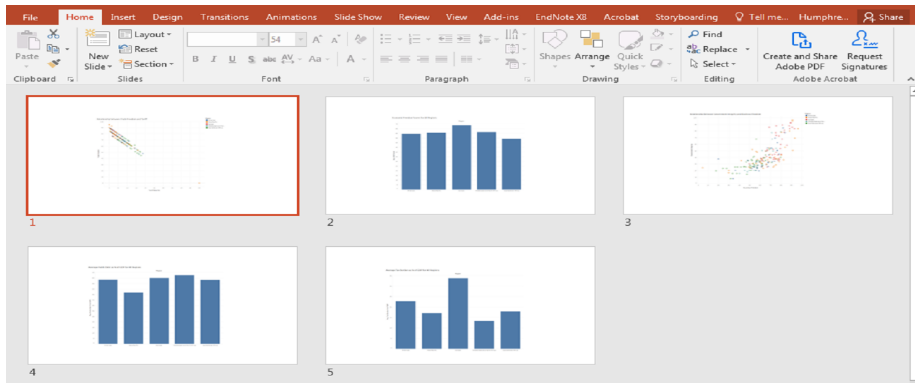
(b) Sequencing the visualisation story pieces in a presentation tool.

(c) Recording the visual data story presentation using a screen recording tool.

Figure 1: Creating a visual data story using 3 different tools for the different stages.

this task. All of the visualisations serve as story pieces and form the basis of the intended visual data story.

At this point, Matt is ready to sequence the visualisations currently in the presentation software. Taking his target audience into consideration, Matt attempts to build a coherent plot by sequencing the selected story pieces in an order he believes to be logical and easy to understand. However, because Matt lacks an empirical method to verify that the sequenced story pieces are more or less comprehensible, he simply moves the story pieces around till he is satisfied with his current sequence based on intuition, as shown in Figure 1b.

Matt also intends to communicate transparency and integrity of ethics to his target audience, so he adds additional slides that capture data acquisition and data transformation information. With his presentation slides ready, Matt is ready to record a presentation for his target audience.

Finally, Matt starts a screen recording software and records his visual data story on the 2019 economic freedom of regions using his presentation slides containing his ordered story pieces and reference notes (Figure 1c).

The above scenario is typical of a visual data story creation process. The process demands a substantial degree of time and effort and a combination of tools. Also, current tools do not incorporate models to support the recommendation of logical story pieces sequences. Moreover, errors made at any stage or component (e.g., in the visualisation authoring software) are propagated to subsequent stages (e.g., presentation software). This is tedious to correct since the errors have to be corrected in the source authoring tool used at the stage of the visual data story creation process before updating subsequent stages.

We aim to improve this workflow by consolidating the different phases of the visual data story creation process in a single tool to enable flexible switching between the different stages of the visual data story creation process and help data analysts like Matt produce coherent and logically sequenced visual data stories. We propose Gravity, a tool that supports the authoring of visualisation story pieces, the recommendation of logical sequences of story pieces, creation of basic presentation slide decks with reference notes, and the recording of the

7

visual data story presentation.

3. Related Work

In this section, we review prior literature on narrative visualisations and tools for authoring visual data stories in the information visualisation field.

3.1. Narrative Visualisation and Visual Data Storytelling

The field of information visualisation is witnessing a growth in research interest in investigating approaches for creating effective visual data stories [16] and criteria for evaluating them [21]. Beginning from the earliest concepts of narrative visualisation proposing the advantages that ensue from the combination of information visualisation and storytelling mechanisms [3, 4], there has been much recent work characterising various perspectives in translating insights into visual data stories [22, 2, 13, 7, 23, 24, 25]. Similarly, there has been an argument for the adoption of narrative visualisations [13, 8, 2] to fulfil the communicative goal of information visualisation.

Segel and Heer [7] presented a framework for narrative visualisations by characterising 58 narrative visualisation artefacts from a variety of sources into 7 genres: magazine style, annotated charts, infographics, flow charts, comic strips, slideshows, and data videos. The scope of our work is limited to the slideshows and data videos genre of the narrative visualisations framework. Segel and Heer's framework also categorises narrative visualisation on a continuum of author and reader-driven strategies.

Other studies have focused on the storytelling scope and creation models [13, 8]. Lee et al. [13] posit the visual data story creation process; the visualisation of specific data as story pieces, the sequencing of these story pieces in a logical order, and the presentation of the sequenced denouement to a target audience. Their work also highlights 4 common visual story authoring settings including live presentations and recorded videos.

Researchers have also examined the effects of sequencing and consistency of visualisation artefacts, i.e.,, story pieces on the comprehension and memorability

8

of narrative visualisations [26, 27, 17, 12, 11]. The remarkably similar results emerging from these studies reveal that effective visualisation sequences significantly improves a target audience perception, comprehension and memorability of narrative visualisations than less optimal sequences, indicating the need for authoring tools to incorporate optimal sequence recommenders amongst other features in tool development.

More recently, Amini et al. [21] proposed, inter alia, engagement, memorability and comprehension as criteria for evaluating visual data stories. The result of a study [24] from 3 web-based experiment exploring engagement in visual data stories showed that the presence of introductory narratives does not necessarily improve user engagement in data exploration. However, author-driven narratives do significantly aid user comprehension and short-term memorability of facts and value messages of information visualisation, and also raise issues of subjectivity and bias [9].

While the aforementioned work has provided valuable insight into the models and structural processes of visual data stories, there has been little work consolidating the various features. Our work bridges this gap and enables users to author interactive visualisations and visual data stories, thus fulfilling both the exploratory and communicative goals of information visualisation.

3.2. Visual Data Story Authoring Tools

There has been an increased adoption of data-driven stories as a medium for conveying data to a target audience in varied domains. This has motivated both academic and industrial researchers to investigate different approaches in developing tools that support the authoring process of data-driven stories [16, 28].

To ease the creation of narratives based on spatiotemporal data for analysts, Eccles et al. [29] proposed GeoTime based on core features designed to provide a higher level of abstraction that was unavailable in previous visualisation systems. GeoTime introduces an automated pattern finding component that searches through the data for behaviours and relationships among events and

entities. These patterns are then highlighted by the annotation component of the system. The system also provides a text editing component for analysts to author their visual data stories and provide explanations of events in time and space.

While GeoTime is specialised and focused on spatiotemporal narratives based on spatiotemporal data, Gravity is a more generic system for authoring visual data stories using any data that can be visualised using common visualisaton techniques.

Ellipsis [1] introduces a model of storytelling abstractions and supports the creation of narrative visualisations using its graphical user interface. The Ellipsis system supports the import of already created visualisations and the use of textual and graphical annotations to create multiple story scenes. However, Ellipsis is focused primarily on the communicative aspect of information visualisation and does not provide support for exploration. On the other hand, Gravity supports both the exploratory and communicative goals of information visualisation. Gravity supports the creation of interactive visualisations and direct import of independent, externally created visualisation specifications. In addition, Gravity consolidates the three processes of visual data story creation; creation of interactive visualisations, sequencing of created visualisations, and the presentation of the logically sequenced visualisations.

SketchStory [30] is a data-driven digital whiteboard that merges infographics with animation. SketchStory enables presenters to quickly create expressive visualisations using pen and touch in real-time presentation sessions. It is built on the premise of synchronous presentation, while the presenter addresses the audience in real-time. By contrast, Gravity's aim is to enable the creation of visual data stories in the form of videos that could be presented asynchronously.

Leveraging the familiar medium of comics, Kim et al. [31] introduced Data-Toon, a data comic storyboarding tool that integrates the analysis and presentation of networks with pen and touch interaction. Using DataToon, analysts can rapidly author visualisations of network panels and annotate them to create a narrative. While DataToon focuses on the visualisation of networks in the comic strip genre of the narrative visualisation framework [7], Gravity addresses the

10

slideshows and data videos genre.

Motivated by a study of how designers produce visualisations in practice,
Bigelow et al. [20] presented Hanpuku, a tool to bridge the gap between illustration tools and visualisation tools. Hanpuku provides an interface that merges D3 scripts into Adobe Illustrator documents to support visualisation iteration. To support designers' workflow, Data Illustrator [32] extends interaction techniques in vector design tools for direct manipulation of visualisations and lazy data binding for visualisation authoring. Moreover, ChartAccent [33] facilitates the communication goal of information visualisation by aiding analysts to easily augment visualisations via a palette of annotation interactions that generate manual and data-driven annotations. Tableau's [34] story point feature allows analysts to present a set of interactive dashboard sheets or views with annotations and filters. Similar to Ellipsis [1], these tools are well-suited for authoring presentation-oriented visual data story pieces. Nonetheless, they do not provide support for optimally sequencing visual data story pieces to produce coherent visual data stories, and also lacks functionality for voice narration and recording of the presentation.

DataClips [16] was introduced as a tool for creating short data-driven motion graphics. DataClips aims to lower the entry threshold to creating short animated clips for non-experts; it emphasises pictographs and has a large library of clips. CLUE [35] is a model that integrates data exploration and presentation based on provenance data captured during the exploration phase. Using CLUE, analysts can author narrative visualisations based on the history of data exploration. However, both DataClips and CLUE do not provide a mechanism for the creation of logical visualisation sequences - a crucial feature for the creation of coherent visual data stories.

4. Design Methodology

4.1. Design Approach

Previous work in information visualisation focused largely on the exploratory aspect of the visualisation process [8]. However, it provides little support in facilitating a holistic design for creating visual data stories and flexible switching between the different components making up the visual data story creation process [36].

Moreover, previous studies [9, 37] have shown that visual data stories significantly aid the understanding of information visualisations to a target audience. Motivated by these works, our objective is to design a system for creating visualisations serving as story pieces in a visual data story, sequencing these visualisations in a logical manner that promotes understanding and reducing cognitive load on the user, and providing support for the presentation of the visual data story. Using these goals as a starting point, we identify 6 high-level tasks that the system should fulfil. In addition, we utilise the *multi-level typology of abstract visualisation tasks* [38] to specify the task descriptions for each task.

The multi-level typology of abstract visualisation task offers abstract rather than domain-specific description of tasks. This characteristic of the typology aids the analysis necessary for the creation of new visual designs or system features, by informing the conversion of domain-specific problems into abstract tasks, and the qualitative evaluation of visualisation (or tool) usage. To specify the task descriptions for the system, we captured the domain-specific (i.e., visual data storytelling) requirements principally from the narrative visualisation literature, e.g., Lee et al. [13], Kosara and Mackinlay [8], Segel and Heer [7], Chevalier et al. [39], and also from the authors' experience in authoring visual data stories. The abstract task analysis is a result of the consensus of two of the authors.

Although the abstract high-level task analysis described in this work is applied to and supports the development of our system, it can also be utilised by other information visualisation researchers and tool developers as fundamental design requirements to support the development of visual data story authoring tools. A

summary of the key tasks and example usage is shown in Table 1.

Table 1: Summary of Tasks and Example Usage

Task No	Task Description	Example Usage
1	Create visualisations	Provide visualisations as story pieces from raw data
2	Create and visualise sequences	Construct meaningfully ordered sequence of story pieces
3	Create insight notes for visualisations	Create notes or annotations associated with findings
4	Provide overview of all visualisations	Enable insight discovery and modification of prior work sessions
5	Inspect individual visualisations	Look-up and inspect fully interactive visualisations for more details
6	Create visual data story	Present and record logically sequenced visual data story for a target audience

4.2. Tasks Descriptions

4.2.1. Task 1: Create Visualisations

Firstly, Gravity should support the task of ***producing*** visualisations from raw data, either as a part of the exploration process or the first phase of the visual data story creation process. These individual interactive visualisations serve as story pieces in the overall visual data story. This is achieved by inputting the raw data, ***selecting*** certain attributes of the data, optionally ***aggregating*** the data, and ***encoding*** the selected attributes into appropriate visualisations. The visualisations can then be ***recorded*** (i.e., exported or saved) as visualisation artefacts for ***presentation*** purposes outside the system or as input to the visualisation sequence component.

For example, from section 2, Matt can explore the *2019 economic freedom* dataset using Gravity. He can create all 5 interactive visualisations that can be exported as individual visualisation artefacts or as story pieces to be used in his visual data story.

13

4.2.2. Task 2: Create and Visualise Sequences

To **produce** visualisation sequences, our system **encodes** visualisations as graph nodes in a graph-based meta-visualisation. The nodes can be **selected** and **arranged** in a chosen order by creating directed edges between nodes. The system also provides a visualisation similarity model for recommending optimal visualisation sequences. In addition, prior saved sequences can be **imported** and **changed** to more efficient sequences before **recording** (i.e., export) or as input to the presentation component.

For example, in section 2, instead of Matt ordering the visualisation story pieces using just intuition to find a sequence that is both logical and comprehensible to his target audience, he can use Gravity's inbuilt optimal sequence recommender or a combination of the recommender and a manually created sequence.

4.2.3. Task 3: Create Insight Notes for Visualisations

One of the goals of information visualisation is getting insight from data. Documenting and communicating this insight to a target audience during a visual data storytelling session is important. At a minimum, our system should enable an analyst to record and **produce** insight notes during data exploration and visual data storytelling creation process. These notes serve as a useful reference for the analyst during the presentation session.

To add notes to a specific visualisation, the analyst **selects** the target visualisation node and enters a note (**annotation**) that may be insights or key points to be referenced during the presentation, and then **records** the note to the visualisation node.

For example, using Gravity, Matt can add insight notes to each of the visualisation story pieces. These notes will be dynamically bound to the visualisations and their corresponding presentation slides, analogous to reference notes in presentation software.

14

4.2.4. Task 4: Provide Overview of All Visualisations

Providing an overview of prior created visualisations or sequences of visualisations is an important task in supporting collaboration. This task functionality enables an analyst to **discover** insight from the prior work of team members and supports the continuation of prior work by **exploring** and **querying** insight notes, visualisations and sequences of visualisations.

The analyst can **import** a Gravity *gty* file and see an ensemble of the visualisations and visualisation sequences of a prior session. They can **select** and **re-arrange** visualisation sequences, create new visualisations, modify existing visualisations and insight notes (**change**), and **produce** an input to the presentation component or **record** the current state of the session.

For example, suppose Matt is a member of a team, he can access, review, and continue prior work of team members on the *economic freedom* dataset including the visualisation story pieces, sequences, and insight notes from a single file in Gravity.

4.2.5. Task 5: Inspect Individual Visualisations

An analyst can view fully interactive mini-previews of visualisations encoded as graph nodes for sequencing purposes. This functionality aids comparison and enables analysts to inspect visualisations while adding insight notes without leaving their current view.

However, to access a mini-preview of a visualisation (i.e., to **discover** the underlying visualisation **encoded** as a graph node), the analyst can **look-up** and **identify** a graph node, **select** the node to render the fully interactive mini-preview and then select the inspect button to render the visualisation in full resolution for inspection in the visualisation component.

For example, in inspecting a visualisation to get more details on certain data points before adding insight notes in section 2, using Gravity, Matt can access and interact with an interactive visualisation while adding the notes at the same time, without leaving his current view.

15

4.2.6. Task 6: Create Visual Data Story

The final stage of the visual data story creation process is the actual presentation of the data and interactive visualisations to a (potential) target audience. Consequently, the system should allow an analyst to **produce** a recorded presentation. This is achieved by the analyst filling an author and ethical consideration form within the system. A basic presentation slide deck is automatically created by the system containing an ordered sequence of fully interactive visualisations (after the introduction slide), with each visualisation slide containing its corresponding insight note. The analyst can then **present** the visual data story by **navigating** through the slides. The system **records** the slideshow and voice narration of the analyst and provides functionality for previewing and exporting the recorded visual data story video.

For example, Matt can generate a basic presentation slide deck from his visualisation story pieces. He can also add an introduction and data acquisition/transformation slides. Moreover, he can present and record the presentation of his visual data story.

5. System Design

Our abstract high-level tasks described in section 4 inherently support the visual data story creation process - creating story pieces, i.e., visualisations, establishing a logical flow between story pieces, and presenting the visual data story [13, 36].

5.1. System Components in Gravity

The main components of the Gravity system are shown in Figure 2. The Visualisation View is where new interactive visualisations are created (**Task 1**). It also supports the inspection of mini-previews of visualisations encoded as graph nodes in the Sequence View in full scale (**Task 5**). The Sequence View facilitates the creation and meta-visualisation of visualisation sequences (**Task 2**) encoded as nodes and edges, and provides an overview of all visualisations

16

created (**Task 4**). Insight notes for each visualisation are also added in the Sequence View (**Task 3**). The Presentation View supports the creation of a basic presentation slide deck with insight notes, and the recording and export of the presentation as a video (**Task 6**).
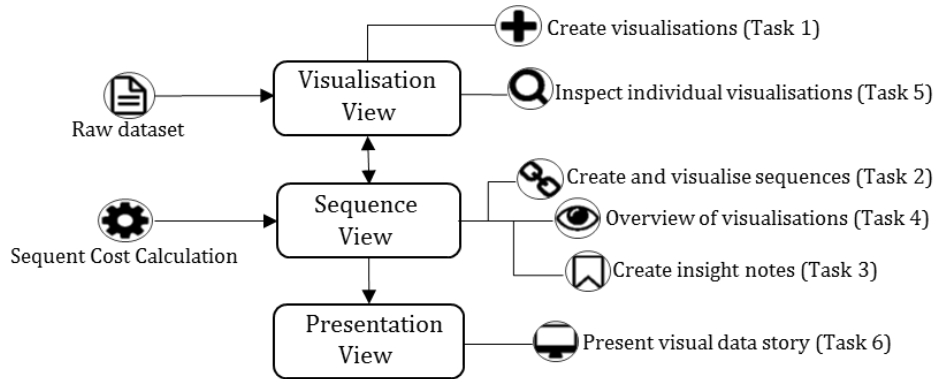

Figure 2: An overview of the Gravity system, showing the three main components.

5.2. Visualisation View

The Visualisation View, shown in Figure 3, provides functionalities to inspect prior created visualisations in full scale and the creation of new ones. To create a visualisation, an analyst inputs data into the system using the file input control (3a), sets the visualisation encodings and aggregation using a set of drop-down menus (3b, 3c), before selecting an appropriate visualisation template (3h) to create the visualisation (3d). A visualisation title is then entered for the created visualisation (1e). The visualisation template panel (3h) contains frequently used visualisation types, namely scatterplot, bar chart, stacked bar chart, line chart, area chart, histogram, box-plot and punch-card plot. When any of the visualisation thumbnails is moused over, the name of the visualisation is shown in a pop-over along with suggested valid data attributes type (3g) **(Task 1)**.

The visualisations are encoded using Vega-lite specifications [40] and can be exported as SVG or PNG files (3f). To use the created visualisation as an input graph node for sequencing purposes in the Sequence View and for later

17

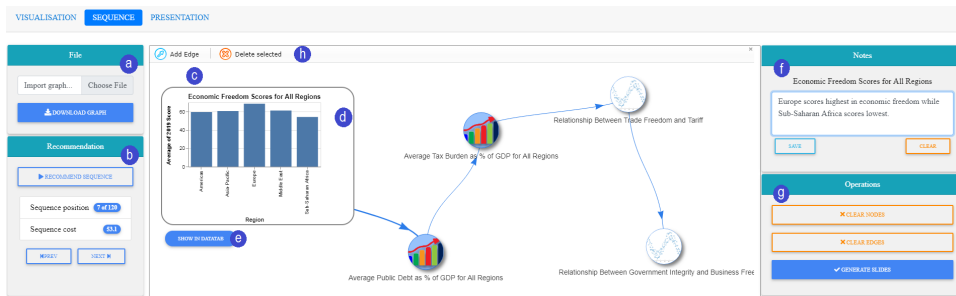Figure 3: The Gravity system's interface: visualisation view.

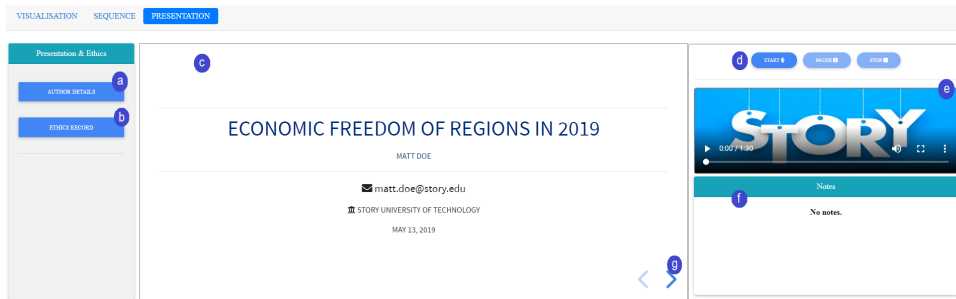

Figure 4: The Gravity system's interface: sequence view.

Figure 5: The Gravity system's interface: presentation view.

presentation, the analyst clicks on the *Create Node* button (3i), which then encodes the visualisation as a graph node in the Sequence View (4c).

5.3. Sequence View

The Sequence View, shown in Figure 4, provides an overview and meta-visualisation of created visualisations and allows for both manual visualisation

18

sequencing and automatic sequence recommendation. The graph canvas (4c)
can be populated either by using the *Create Node* button in the Visualisation
View (3i) or by importing a previously saved *gty* file using (4a). Individual
visualisations are encoded as graph nodes and directed edges connecting them
show their sequence order. The icon of a graph node represents what type of
visualisation it is encoding, e.g., bar chart, using the same thumbnail icons as in
the visualisation template panel (3h). The title of the graph node is the same as
the title of the visualisation it is encoding for identification purposes.

An analyst can add insight notes to each graph node (for referencing purposes
in the presentation view, 5f) by selecting the graph node and filling the textbox
(4f); the title of the textbox is bound to the title of the graph node for easy
identification **(Task 3)**.

Mini-previews (4d) of the visualisations are provided to enable analysts to
see the underlying visualisations encoded as graph nodes without having to leave
the Sequence View. This facilitates comparison and addition of insight notes.
However, to inspect the mini-preview visualisation in full scale, the analyst clicks
on the *Inspect* button (4e) which then renders the visualisation in full scale in
the Visualisation View **(Task 5)**. This allows the easy update of graph nodes.

An analyst can manually create a visualisation sequence by selecting the *Add
Edge* button (4h), clicking on a visualisation node and then dragging the edge
to another visualisation node. This process is repeated until all graph nodes are
connected in a preferred sequence. However, to use the system recommended
logical sequences, the analyst can click on the *Recommend Sequence* button in
the recommendation panel (4b). The system computes the recommendation
using the Graphscape model [11] in worker threads and displays the results,
i.e., recommended logical sequences. The *Sequence position* shows the currently
selected sequence out of a total number of recommendation, while the *Sequence
cost* is a summary outcome from the Graphscape model; effective sequences
have smaller costs. For example, Figure 4b shows sequence #7 out of 120
recommendations, with a sequence cost of 53.1 (by comparison, the least optimal
sequence has a cost of 91.4). The analyst can navigate through the recommended

19

sequences before selecting an appropriate one **(Task 2)**.

The *Clear Nodes* and *Clear Edges* buttons in the operations panel (4g) are used for clearing all nodes and edges while the *Generate Slide* button allows the analyst to output the selected sequence to the Presentation View. In addition, the current state of the graph can be exported as a *gty* file (4a) and re-imported at a later time for collaboration purposes **(Task 4)**.

5.4. Presentation View

The Presentation View, shown in Figure 5, supports the creation and presentation of a visual data story. The analyst clicks on the *Author Details* button (5a) which opens up a modal form to fill in information about the visual data story presenter or analyst; the analyst also completes the ethical consideration form by clicking on the *Ethics Record* button (5b). The purpose of the ethical consideration form (which automatically fills the content of the penultimate slide) is to promote ethics in visual data storytelling as proposed by previous works [41, 9, 25]. Using the ethical consideration form, the analyst can show the data source and the data transformations employed during data exploration.

The ordered sequence selected (*Generate Slides* button, 4g) containing fully interactive visualisations and insight notes is contained as part of the slide deck. The analyst then presents the visual data story using the prepared presentation slides and the recording functionality which records both the presentation slides and voice narration. The slide deck (5c) can be navigated using the controls (5g). The notes panel (5f) shows the corresponding notes created for each visualisation in the Sequence View (4f) as the analyst navigates through the presentation slides.

To record the presentation, the analyst clicks on the *Start* button (5d) and gives microphone access to the system. Lastly, the recorded video can be previewed and exported in 5e **(Task 6)**.

5.5. Visualisation Similarity and Sequence Recommender

Prior works have shown that the order or sequence in which visualisations are shown to a target audience affects comprehension and memorability; optimal

sequences decrease cognitive load and promote better comprehension and recall of the value message of the visual data story while less optimal sequences impede user comprehension and recall [17, 12]. Hence, the introduction of an effective sequence recommender in Gravity.

The Graphscape Model is a directed graph model for reasoning about visualisation similarity and sequences[11]. It represents visualisation specifications as graph nodes and transformation between nodes as directed edges. The model proposes an estimated cost of difficulty of interpreting a target visualisation given a source visualisation. The model derives transition costs for a directed graph of visualisations by a partial ordering of edit operations consisting of translation between mark types (i.e., geometric primitives), transformation functions and encodings between pairs of visualisations. Within the model, mark type transitions are the least expensive edit operation (since simply changing a mark type holds all data constant), followed by data transformation (which retains the same underlying data but changes the summary level of the data). The most expensive edit operation is visual encoding (since this essentially changes the data fields being visualised). In short, *mark type* < *data transformation* < *visual encoding operations*. The sum of all costs in one category (e.g., *data transformation*) has to be less than any single operation in a more expensive category (e.g., *visual encoding operations*).

The Graphscape model is a directed graph, $Graphscape = (V, E, D)$, where nodes $v \in V$ are Vega-lite visualisation specifications [40] and edges $e \in E$ are edit operations that transform a source visualisation to a target visualisation [11]. The model consists of 3 components that are combined together to generate the final sequence costs for a set of visualisations. The components are transition costs, filter sequence costs, and global weighting.

5.5.1. Transition Costs

Each edge e in the model has edit cost $w(e)$ that is determined based on the edit operations ranking, i.e., *mark type* < *data transformation* < *encoding operations*. Generally, transition cost $T(u, v)$ between a source visualisation u

21

and a target visualisation v is defined as the sum of edge weights along the shortest path between them:

$$T(u, v) = \Sigma_{e \in ShortestPath(u,v,w)} w(e)$$

5.5.2. Filter Sequence Costs

Filter operation focuses on the *equality* predicates which are common in visual analysis. The model compares data values within *equality* predicates and rewards sorted values. Given an input sequence S of visualisations, all recurring filtered fields within the sequence are identified and a set V consisting of sequences of predicate graph nodes values v_i is extracted. The filter cost term F is given by:

$$F(S) = 1 - \frac{1}{|V|} \sum_{v \in V} \frac{|\sum_{i=2}^{|v|} d(v_{i-1}, v_i) + 0.1|}{|V| - 1 + 0.1}$$

$$d(v_a, v_b) = \frac{v_a - v_b}{|v_a - v_b|} \ \ if \ v_a \neq v_b, 0 \ otherwise$$

For each filtered field, the ascending and descending value changes are scored as $+1$ and -1, respectively. The absolute value of the sum of scores for each field is calculated and normalised by the number of filter changes. The additional 0.1 biases the score in favour of ascending order (alphabetical order for categorical data and chronological order for dates). The average of all filter scores is taken and subtracted from one to convert the score to a cost.

5.5.3. Global Weighting

The Graphscape model includes a global weighting term W to reward sequences that order visualisations into consistent parallel subsequences. It defines P as a repeated, non-overlapping subsequence of identical transitions.

The global weighting term W is determined by the fraction of the sequence S covered by the pattern P that induces maximal coverage over S:

$$W(S) = 1 - max_P \frac{count(P, S) \cdot |P|}{|S|}$$

22

The $W(S)$ term decreases the overall sequence cost by a multiplicative factor that reflects the number of consistent subsequence transitions. If there is no repeating pattern P found, W is set to 1 and the sequence cost remains the same.

530 *5.5.4. The Graphscape Sequence Cost Function*

Combining all of the above components together, i.e., transition costs, filter sequence costs, and global weighting costs, the Graphscape cost function for a visualisation sequence S is given as:

$$Cost(S) = W(S) \cdot \left(F(S) + \sum_{i=1}^{|S|} T(S_{i-1}, S_i) \right)$$

To account for the cost of interpreting an initial visualisation within a
535 sequence, each input sequence S is treated as having an initial entry S_0 consisting of a null visualisation specification in which no encoding fields are specified.

Sequences generated by the Graphscape model have been evaluated and shown to be effective [11, 22].

For Gravity, we adapt and operationalise the Graphscape model for recom-
540 mending the most optimal sequences of visualisations.

In addition, we provide a meta-visualisation of the graph nodes and edges, in order to support the visualisation of the recommended sequences. We encode visualisation sequences as nodes and directed edges in a graph canvas. The nodes are visualisation specifications and the edges between nodes show a logical
545 connection between the nodes. Our approach also facilitates the creation of manual sequences within the graph canvas. We assumed the meta-visualisation approach for two reasons; firstly, we aimed to concretise the abstract Graphscape model to help users intuitively understand and interact with the (recommended) visualisation sequences in addition to the model recommendation results; and
550 secondly, meta-visualisation displays have been shown to be effective in supporting visual understanding when there are numerous visualisations to be analysed together, in this case, multiple visualisation story pieces in a single sequence

view [18, 42].

Topological sorting of a directed graph is a linear ordering of its nodes such that for every directed edge UV from node U to V, U comes before V in the ordering [43]. In the context of this work, the nodes of the graph represent visualisations to be presented and the edges represent the constraint that one visualisation must be presented before another in a linear order.

In Gravity, we implement a topological sorting algorithm to sort and render a selected visualisation sequence in a linear order for presentation purposes in the presentation view.

5.6. A Note on Implementation

Gravity is a web-based prototype application developed using the Vuejs[5] framework in a component-based approach. It is developed completely as a client-side tool with no server installation or maintenance. Gravity currently accepts raw data files in a comma-separated values (CSV) format, where each row represents a data record and each column a data attribute. The visualisations are created using Vega-lite [40] specifications; a high-level grammar of interactive graphics that makes it relatively easy to generate web-based interactive visualisations, and is expressed in concise JSON formats.

In creating the meta-visualisation as a directed graph of visualisation nodes and edges, the generated Vega-lite visualisation specifications are passed (with unique IDs) as JSON objects from the visualisation view to the sequence view.

The recommender sequence properties are calculated using the Graphscape model [11] implemented in Javascript. The model computes the encoding distance between pairs of visualisations by calculating the sum of edges based on the shortest path traversal using the breadth-first search algorithm. The breadth-first search is quite expensive to be run in a browser, hence we delegate the sequence recommendation computation to separate worker threads using the client's CPU to prevent the user interface from becoming unresponsive. The

---

[5]`https://vuejs.org`

recommended sequence of visualisations nodes and connected edges are passed from the sequence view to the presentation view.

The presentation of the logical sequences of visualisations is recorded in the presentation view. Currently, there is no support for the recording of HTML elements (except for canvas and media elements). To address this challenge, the parent HTML node for rendering the visualisations are read from the DOM into a canvas element at rapid intervals. The mediastream recording API is then used to record the canvas, as well as capturing the audio, which are then merged and converted to videos in the Webm[6] format.

6. Evaluation

We carried out a user study to evaluate the visual data story creation process of Gravity, guided by the evaluation methods proposed by Amini et al. [21] and Ren et al. [44].

6.1. Participants and Apparatus

We recruited 10 representative participants (5 male, 5 female) including 6 expert visualisation designers and analyst and 4 non-experts with basic visualisation experience. Two of the experts were recruited from a Data Consulting firm based in Melbourne, Australia, three are postgraduate students doing research in data science and visualisation, and one is a university research engineer. The expert participants produce presentations based on data in the course of their work. The four remaining non-experts with basic visualisation experience have varied backgrounds in arts & humanities, engineering and business. Participants were between 18 and 40 years old.

We contacted participants via advertisements on university noticeboards and email announcement. Participants were rewarded with a $20 shopping vouchers for completing the study. We did not record any conflicts of interest.

---

[6]https://www.webmproject.org/

The hardware apparatus provided to participants were a desktop computer running Microsoft Windows 7 OS, a 24-inch monitor and standard peripheral devices - keyboard and mouse.

610 The objective of the study was to obtain insight into how Gravity supports the creation of logically sequenced visual data stories using a workflow that consolidates the different phases of the visual data story creation process.

We note here that we do not evaluate the effectiveness of the logical sequences, since Kim et al. [11] reported on a comprehensive result of the effectiveness

615 of Graphscape sequences. Rather, we evaluate Gravity, the visual data story authoring system which operationalises and visualises the Graphscape model as part of the system's components, using the visual data story authoring systems evaluation criteria proposed by Amini et al. [21].

6.2. Procedure

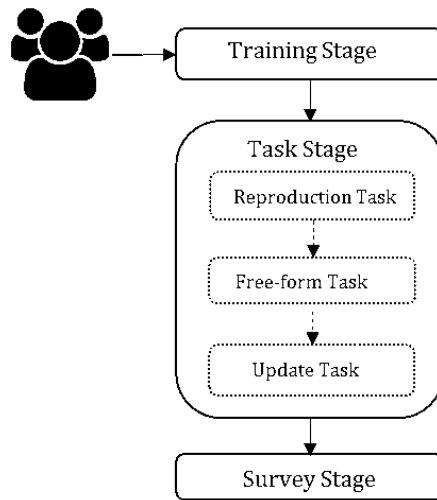620 The user evaluation procedure is summarised in Figure 6.

Figure 6: The user evaluation study procedure.

26

6.2.1. Training Stage

An investigator informed participants of the purpose, benefits and voluntary nature of the evaluation study and requested participants to sign a consent form before the session began.

Participants were introduced to the Gravity system by a researcher who described the system features and interactions. Participants then watched tutorial videos for 6 minutes to reinforce what has been explained to them. They were encouraged to ask any questions before starting the main evaluation tasks.

6.2.2. Task Stage

The tasks were designed such that participants will be able to interact with the main features of Gravity, and also to capture the evaluation criteria for evaluating visual data storytelling tools proposed by Amini et al. [21] and Ren et al. [44]. There were 3 main task types; reproduction, free-form and update tasks. Participants were handed a task sheet containing the tasks descriptions before they proceeded to complete the tasks.

Reproduction Task. The focus of a reproduction study is the usability of a system and the learnability of its features. In reproduction studies, participants are typically asked to complete tasks that reproduce a copy of one or more (visualisation) designs. The reproduction study evaluation method is commonly used in the evaluation of visualisation tools and has been employed in evaluating tools such as Data Illustrator [32], and Lyra [45]. Reproduction studies help to assess if participants can produce the desired designs, and if not, what the main barriers might be [44]. Hence our evaluation includes a reproduction task.

To motivate this task, participants were asked to decide on a potential audience who will be the recipient of their created visual data story. Participants were then given 5 sheets of paper containing 5 visualisations of the open source cars data [46] and asked to reproduce the visualisations before saving them as graph nodes. Participants had to add a minimum of one insight note to each graph node and then create a visualisation sequence either manually or using the

27

system's sequence recommender. They also developed a narrative they would like to communicate to their audience. Lastly, participants completed the author detail and ethical consideration forms, delivered a visual data story based on the created visualisation sequence and then exported the video. The result of this task is a complete narrative sequence, with voice-over narration.

Free-form Task. In a free-form study, participants are typically asked to complete tasks on their own using the system. Free-form studies focus on assessing whether or not participants can produce (visualisation) designs of their own using the new authoring system without significant support [44]. The free-form study method has been employed in evaluating visualisation tools such as Ellipsis [1] and DataClips [16]. Free-form studies also reflect the learnability and usability of a system, as participants go through the process of executing tasks on their own with minimal guidance. In addition, a free-form study can be preceded by a reproduction study, which serves as an additional tutorial, as is the case in our evaluation of the Gravity system.

For the free-form task, participants decided on a potential target audience for their presentation as in the previous task, and developed a narrative they would like to communicate to their audience. They were asked to freely create their own visualisations (minimum of 3), sequences, and visual data story based on any of the 5 sample open source datasets available within the system. The sample datasets include the following: Melbourne temperature [47], iris flower [46], Seattle weather [46], IMDB movies [46], and cars [46]. Each sample data had a description so that participants can have some understanding of the data. The result of this task is a complete narrative sequence, with voice-over narration.

Update Task. The goal of the update task is to capture the asynchronous collaboration support of the Gravity system.

For this task, participants were put in a scenario - they were part of an analytics team and had to continue working on an existing *gty* file based on the work that was left off by a team member. They were to review the prior created visualisations and visualisation sequence, make addition, modification

28

and updates before saving the state of the session as an updated *gty* file. The updates participants made to the existing *gty* file include modifying y-axis, titles, and insight notes of visualisations. They also authored new visualisations and their corresponding insight notes, and updated the visualisation sequence. The result of this task is a modified and updated *gty* file.

6.2.3. Survey Stage

At the end of the study, participants completed a survey. The survey captured demographics information and evaluation criteria such as usefulness, satisfaction, and ease of use, using a 5-point Likert scale. Investigators encouraged participants to provide additional feedback and comments in the spaces provided.

6.3. Results and Discussion

The task questions (attached in the supplementary materials) were designed to cover the abstract high-level tasks 1 - 6 discussed in section 4 to consolidate the visual data story creation process. Moreover, the evaluation study aimed to collect insights on how the Gravity system supports users in creating visual data stories.

Participants were successful in carrying out the evaluation tasks without significant guidance. Participants spent an average of 19 minutes to complete the reproduction task, 13 minutes to complete the free-form task, and 9 minutes to complete the update task. The survey results grouped by different criteria (as proposed by Amini et al. [21] and Ren et al. [44]) is summarised in Figure 7. Overall, the results for the different criteria are positive (more than 85% on average). A detailed table with 37 survey questions and responses spread over different criteria is shown in Table 3 (Appendix A).

6.3.1. Evaluation Categories

Ease of Use. Although the Gravity system is a research prototype with a focus on the most important features, participants agreed on the ease of use of the system. This is reflected in their response to the questions. They found the system easy to use for creating visualisations, visualisation sequences and presenting their
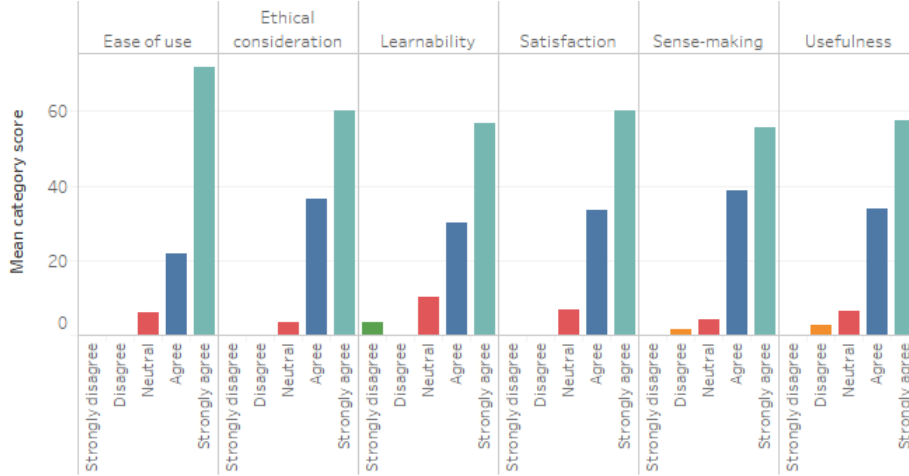
Figure 7: Summary result of the user evaluation study.

visual data stories. For instance, an expert participant P4 commented, "It is user friendly and interactive".

Learnability. The result shows that most participants learned to use the tool quickly within a short training period. However, participants also agreed that they would have loved to receive further instructions. We note that this can be linked to participants' lack of familiarity with the task datasets since an understanding of the social context of the data is important in creating visual data stories. For instance, a non-expert participant P1 commented, "The tool is very impressive...my only setback was trying to understand the sample data".

Usefulness. Participants found the meta-visualisation using nodes and directed edges helpful in creating and understanding visualisation sequences. In addition, they found the insight note and other functionalities useful for recording their understanding of the data and reference during presentation. Overall, participants found the system useful for creating visual data stories.

Sense-making. The evaluation results show that the system supports sense-making from an author perspective. We hypothesise that the author sense-

30

making success may likely cascade into sense-making for a target audience, although another study will be needed to verify this. Participants not only understood the visualisations that they created, but they also agreed that the system aided them in articulating their understanding to their recipient audience.

Satisfaction. As shown in Figure 7, participants responded positively to questions on user satisfaction. They concurred that they will adopt the system for visualisation tasks in future projects and would recommend it to a friend.

Ethical Consideration. Participants found the ethical consideration form in the system an important aspect of ethics in the visual data story creation process. They agreed that the presence of ethical records will promote author credibility and a target audience trust in a presented visual data story.

6.3.2. Discussion

Successes. Participants successfully created visual data stories without significant assistance from the investigators. They authored 2 visual data stories and followed the visual data story creation process; they created individual visualisations which served as story pieces, sequenced these story pieces in a logical order, added insight notes, completed author details and ethical consideration forms, and gave and exported their presentations (with voice-over narration) using the basic presentation slides deck. In short, the system satisfied the 6 abstract high-level tasks described in section 4 - the rationale of the systems' design requirements, which in turn supported participants in authoring logically sequenced visual data stories.

Participants appreciated the system's sequence recommender, especially being able to navigate through the recommendations using the meta-visualisation and sequence cost as guides. All participants used this feature at least once during their evaluation session. Participants selection of the recommended sequences was systematic. Almost all participants selected the recommended sequence based on the displayed sequence cost and how they would like to communicate their intended narrative. Participants also manually created sequences by creating

31

new edges between graph nodes. Three participants stated that the sequence recommendation and linking of the presentation slides were their favourite feature of the system, and was useful in helping them develop their narrative.

Participants also found the meta-visualisation of visualisation sequences useful. They agreed that the meta-visualisation using a graph representation aided them in understanding the system recommended sequences and in creating their own manual sequence.

One participant reported that the support of *gty* files (import and export) is important for collaboration, enabling them to share visualisation artefacts with others. Another participant commended the fully interactive mini-preview visualisations stating, "[mini-preview] helps a lot when I have to add notes to the node". Moreover, participants also welcomed the Gravity system as a user-friendly and interactive web-based system; "I like the fact that the tool is web-based, so one does not need to install…" noted a participant. These successes show that the system fulfils the higher level analysis tasks that were the basis for its design.

Current Limitations. Although participants' response is mostly positive, they also noted potential areas for improvement in the Gravity system. The current implementation of the system allows manual updates of graph nodes (e.g., modification of encodings of the underlying visualisations) in the graph canvas; this results in the duplication of the intended node - the system keeps the old node and creates a newly updated node, and users have to manually delete the old node. Participants suggested that the next version of the system should support the automatic update of nodes without duplication.

Also, the current implementation of the system only shows the sequence cost for system recommended sequences but does not calculate sequence cost for manually created sequences. As one participant notes, "Telling me my manual [sequence] cost would be excellent." These observations show the need for further iteration of the system.

6.4. Threats to Validity

We took considerations to reduce threats to validity and their impact in the user evaluation study. However, there are threats that may have affected the results and conclusion. We discuss these threats and controls applied to them and summarise these in Table 2.

Table 2: Summary of threats to validity.

Threat	Category	Status
Instrumentation	Internal	Addressed
Maturation	Internal	Addressed
Population sample	External	Partially addressed
Sample size	External	Partially addressed

6.4.1. Internal Validity

Instrumentation. We simplified the tasks protocols and encouraged participants to ask questions during evaluation sessions. However, some participants may have been reluctant to ask questions and may have responded based on their understanding of the tasks. This may have potentially affected participants' experience of the evaluation study. We do note that there were no concerns raised regarding tasks protocols.

Maturation. There is a risk of participants experiencing boredom and lethargy as the evaluation study progressed over time. However, investigators endeavoured to encourage participants. They were awarded a $20 voucher as a reward for their participation. Also, participants were advised that they were under no obligation to complete the evaluation study and could leave whenever they wanted.

6.4.2. External Validity

Population Sample. Our goal was to get a representative sample of information visualisation experts and non-expert participants. We successfully recruited 6 expert participants and 4 non-experts. All of our expert participants author

visual data stories in the course of their work. Two of the expert participants work as consultants in a Data Consulting firm based in Melbourne, Australia, 3 are postgraduate (PhD) students engaged in active research in data science and visualisation, while the remaining one is a university research engineer. The 4 non-expert participants do have basic experience with data visualisation and have varied backgrounds in arts and humanities, engineering, and business.

Sample Size. We recruited a total number of 10 participants (5 male, 5 female) for the user evaluation study. It is probable that the number of participants may have affected the results of the study. However, since the goal of the evaluation study was not the comparison of participants' performance, but rather to get insights on how the system supports users in the creation of visual data stories and possible areas for improvement, we believe the impact of this is limited. Moreover, the number of expert participants is in the accepted range of participant numbers in information visualisation systems evaluations [45, 1, 16, 30, 31].

7. Conclusion and Future Direction

As enthusiasm for visual data stories grows in the industry and academic research, the development of more tools would be integral to the authoring process of logically sequenced visual data stories. Our paper presents Gravity, a system that consolidates the visual data story creation process and embeds and visualises a model for logically sequencing visualisations to support the presentation of coherent narratives. Pairs of logically related visualisations with the least difficulty in interpretation are connected to each other.

Additionally, we present our design rationale, based on a high-level task analysis that the system should satisfy. Results of our user evaluation study show that the system supports users in creating visual data stories based on the high-level tasks and provide insights on how the system aids users during the creation process.

34

The Gravity system is a proof-of-concept prototype to show the viability of supporting the authoring of logically sequenced visual data stories. Gravity currently supports only a limited number of visualisation types. Further improvements to the system including the refinement of the user interface, the addition of more visualisation types and support for more statistical features to carry out different types of analysis may prove to be useful. Supporting these and much more features without making the system too cumbersome to learn and use is a possible avenue for future work in developing visual data story authoring tools.

Furthermore, we aim at investigating possible secondary applications for the current implementation of our meta-visualisation approach; we will seek further auxiliary uses for the graph nodes and directed edges beyond its primary usage for logical sequencing. For instance, expressing data navigation and user interaction using graphs to enable more flexible alteration of visualisation artefacts for visual data story creation is a potential use case. Nonetheless, more work is needed to model and formalise this objective.

8. Acknowledgements

References

[1] A. Satyanarayan, J. Heer, Authoring narrative visualizations with ellipsis, Comput. Graph. Forum 33 (3) (2014) 361–370. doi:10.1111/cgf.12392. URL http://dx.doi.org/10.1111/cgf.12392

[2] K. Ma, I. Liao, J. Frazier, H. Hauser, H. Kostis, Scientific storytelling using visualization, IEEE Computer Graphics and Applications 32 (1) (2012) 12–19. doi:10.1109/MCG.2012.24.

[3] N. Gershon, W. Page, What storytelling can do for information visualization, Commun. ACM 44 (8) (2001) 31–37. doi:10.1145/381641.381653.
URL http://doi.acm.org/10.1145/381641.381653

[4] W. Wojtkowski, W. G. Wojtkowski, Storytelling: its role in information visualization, in: European Systems Science Congress, 2002.

[5] N. L. Stein, V. I. Kissel, Routledge Encyclopedia of Narrative Theory, 1st Edition, Taylor and Francis Group, Oxfordshire, UK, 2005.

[6] A. Graesser, B. Olde, B. Klettke, Narrative Impact: Social and Cognitive Foundation, Mahwah, NJ: Lawrence Erlbaum Associates, 2002, Ch. How does the mind construct and represent stories.

[7] E. Segel, J. Heer, Narrative visualization: Telling stories with data, IEEE Transactions on Visualization and Computer Graphics 16 (6) (2010) 1139–1148. doi:10.1109/TVCG.2010.179.

[8] R. Kosara, J. Mackinlay, Storytelling: The next step for visualization, Computer 46 (5) (2013) 44–50. doi:10.1109/MC.2013.36.

[9] H. O. Obie, C. Chua, I. Avazpour, M. Abdelrazek, J. Grundy, T. Bednarz, A study of the effects of narration on comprehension and memorability of visualisations, Journal of Computer Languages 52 (2019) 113 – 124. doi:https://doi.org/10.1016/j.cola.2019.04.006.
URL http://www.sciencedirect.com/science/article/pii/S1045926X19300059

[10] T. Gnambs, M. Appel, C. Schreiner, T. Richter, M.-B. Isberner, Experiencing narrative worlds: A latent state–trait analysis, Personality and Individual Differences 69 (2014) 187 – 192. doi:https://doi.org/10.1016/j.paid.2014.05.034.
URL http://www.sciencedirect.com/science/article/pii/S0191886914003286

[11] Y. Kim, K. Wongsuphasawat, J. Hullman, J. Heer, Graphscape: A model for automated reasoning about visualization similarity and sequencing, in: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, CHI '17, ACM, New York, NY, USA, 2017, pp. 2628–2638. `doi: 10.1145/3025453.3025866`.
URL `http://doi.acm.org/10.1145/3025453.3025866`

[12] J. Hullman, R. Kosara, H. Lam, Finding a clear path: Structuring strategies for visualization sequences, Comput. Graph. Forum 36 (2017) 365–375.

[13] B. Lee, N. H. Riche, P. Isenberg, S. Carpendale, More than telling a story: Transforming data into visually shared stories, Computer Graphics and Applications, IEEE 35 (5) (2015) 84–90. `doi:10.1109/MCG.2015.99`.

[14] K. Wongsuphasawat, D. Moritz, A. Anand, J. Mackinlay, B. Howe, J. Heer, Voyager: Exploratory analysis via faceted browsing of visualization recommendations, IEEE Transactions on Visualization and Computer Graphics 22 (1) (2016) 649–658. `doi:10.1109/TVCG.2015.2467191`.

[15] M. A. Yalçın, N. Elmqvist, B. B. Bederson, Keshif: Rapid and expressive tabular data exploration for novices, IEEE Transactions on Visualization and Computer Graphics 24 (8) (2018) 2339–2352. `doi:10.1109/TVCG.2017. 2723393`.

[16] F. Amini, N. H. Riche, B. Lee, A. Monroy-Hernandez, P. Irani, Authoring data-driven videos with dataclips, IEEE Transactions on Visualization and Computer Graphics 23 (1) (2017) 501–510. `doi:10.1109/TVCG.2016. 2598647`.

[17] J. Hullman, S. Drucker, N. Henry Riche, B. Lee, D. Fisher, E. Adar, A deeper understanding of sequence in narrative visualization, IEEE Transactions on Visualization and Computer Graphics 19 (12) (2013) 2406–2415. `doi: 10.1109/TVCG.2013.119`.
URL `http://dx.doi.org/10.1109/TVCG.2013.119`

[18] J. Peltonen, Z. Lin, Information retrieval approach to meta-visualization, Machine Learning 99 (2) (2015) 189–229. `doi:10.1007/s10994-014-5464-x`.
URL `https://doi.org/10.1007/s10994-014-5464-x`

[19] A. Bigelow, S. Drucker, D. Fisher, M. Meyer, Reflections on how designers design with data, in: Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces, AVI '14, ACM, 2014, pp. 17–24.

[20] A. Bigelow, S. Drucker, D. Fisher, M. Meyer, Iterating between tools to create and edit visualizations, IEEE Transactions on Visualization and Computer Graphics 23 (1) (2017) 481–490.

[21] F. Amini, M. Brehmer, G. Bolduan, C. Elmer, B. Wiederkehr, Data-Driven Storytelling, 1st Edition, A K Peters/CRC Press, Boca Raton, FL, 2018, Ch. Evaluating Data-Driven Stories and Storytelling Tools.

[22] S. Xu, C. Bryan, K. Li, J. Zhao, K.-L. Ma, Chart constellations: Effective chart summarization for collaborative and multi-user analyses, Computer Graphics Forum 37 (2018) 75–86. `doi:10.1111/cgf.13402`.

[23] F. Amini, N. Henry Riche, B. Lee, C. Hurter, P. Irani, Understanding data videos: Looking at narrative visualization through the cinematography lens, in: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI '15, ACM, New York, NY, USA, 2015, pp. 1459–1468. `doi:10.1145/2702123.2702431`.
URL `http://doi.acm.org/10.1145/2702123.2702431`

[24] J. Boy, F. Detienne, J.-D. Fekete, Storytelling in information visualizations: Does it engage users to explore data?, in: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI '15, ACM, New York, NY, USA, 2015, pp. 1449–1458. `doi:10.1145/2702123.2702452`.
URL `http://doi.acm.org/10.1145/2702123.2702452`

[25] J. Hullman, N. Diakopoulos, Visualization rhetoric: Framing effects in narrative visualization, IEEE Transactions on Visualization and Computer Graphics 17 (12) (2011) 2231–2240. `doi:10.1109/TVCG.2011.255`.

[26] Z. Qu, J. Hullman, Evaluating visualization sets: Trade-offs between local effectiveness and global consistency, in: Proceedings of the Sixth Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization, BELIV '16, ACM, New York, NY, USA, 2016, pp. 44–52. `doi:10.1145/2993901.2993910`.
URL `http://doi.acm.org/10.1145/2993901.2993910`

[27] Z. Qu, J. Hullman, Keeping multiple views consistent: Constraints, validations, and exceptions in visualization authoring, IEEE Transactions on Visualization and Computer Graphics 24 (1) (2018) 468–477. `doi:10.1109/TVCG.2017.2744198`.

[28] P. Isenberg, B. Lee, H. Qu, M. Cordeil, Immersive Visual Data Stories, Springer International Publishing, Cham, 2018, pp. 165–184. `doi:10.1007/978-3-030-01388-2_6`.
URL `https://doi.org/10.1007/978-3-030-01388-2_6`

[29] R. Eccles, T. Kapler, R. Harper, W. Wright, Stories in geotime, in: 2007 IEEE Symposium on Visual Analytics Science and Technology, 2007, pp. 19–26. `doi:10.1109/VAST.2007.4388992`.

[30] B. Lee, R. H. Kazi, G. Smith, Sketchstory: Telling more engaging stories with data through freeform sketching, IEEE Transactions on Visualization and Computer Graphics 19 (12) (2013) 2416–2425. `doi:10.1109/TVCG.2013.191`.
URL `http://dx.doi.org/10.1109/TVCG.2013.191`

[31] N. W. Kim, N. Riche, B. Bach, G. A. Xu, M. Brehmer, K. Hinckley, M. Pahud, H. Xia, M. McGuffin, H. Pfister, Datatoon: Drawing dynamic network comics with pen + touch interaction, in: Proceedings of the 2019

CHI Conference on Human Factors in Computing Systems, ACM, ACM, New York, 2019, pp. 1–12.

URL https://www.microsoft.com/en-us/research/publication/datatoon-drawing-dynamic-network-comics-with-pen-touch-interaction/

[32] Z. Liu, J. Thompson, A. Wilson, M. Dontcheva, J. Delorey, S. Grigg, B. Kerr, J. Stasko, Data illustrator: Augmenting vector design tools with lazy data binding for expressive visualization authoring, in: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, ACM, New York, NY, USA, 2018, pp. 1–13. doi:10.1145/3173574.3173697.
URL http://doi.acm.org/10.1145/3173574.3173697

[33] D. Ren, M. Brehmer, Bongshin Lee, T. Höllerer, E. K. Choe, Chartaccent: Annotation for data-driven storytelling, in: 2017 IEEE Pacific Visualization Symposium (PacificVis), 2017, pp. 230–239. doi:10.1109/PACIFICVIS.2017.8031599.

[34] Tableau, Tableau software, [Online.] Available: https://www.tableau.com/, accessed January, 2019.

[35] S. Gratzl, A. Lex, N. Gehlenborg, N. Cosgrove, M. Streit, From visual exploration to storytelling and back again, Computer Graphics Forum 35 (3) (2016) 491–500.
URL http://dl.acm.org/citation.cfm?id=3071534.3071586

[36] H. O. Obie, C. Chua, I. Avazpour, M. Abdelrazek, J. Grundy, T. Bednarz, A framework for authoring logically ordered visual data stories, in: 2019 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC), 2019, pp. 257–259. doi:10.1109/VLHCC.2019.8818925.

[37] M. Borkin, Z. Bylinskii, N. W. Kim, C. M. Bainbridge, C. S. Yeh, D. Borkin, H. Pfister, A. Oliva, Beyond memorability: Visualization recognition and recall, IEEE Transactions on Visualization and Computer Graphics 22 (2016) 519–528.

[38] M. Brehmer, T. Munzner, A multi-level typology of abstract visualization tasks, IEEE Transactions on Visualization and Computer Graphics 19 (12) (2013) 2376–2385. `doi:10.1109/TVCG.2013.124`.

[39] F. Chevalier, M. Brehmer, G. Bolduan, C. Elmer, B. Wiederkehr, Data-Driven Storytelling, 1st Edition, A K Peters/CRC Press, Boca Raton, FL, 2018, Ch. From Analysis to Communication: Supporting the Lifecycle of a Story.

[40] A. Satyanarayan, D. Moritz, K. Wongsuphasawat, J. Heer, Vega-lite: A grammar of interactive graphics, IEEE Transactions on Visualization and Computer Graphics 23 (1) (2017) 341–350. `doi:10.1109/TVCG.2016.2599030`.

[41] N. Diakopoulos, Data-Driven Storytelling, 1st Edition, A K Peters/CRC Press, Boca Raton, FL, 2018, Ch. Ethics in Data-Driven Visual Storytelling.

[42] H. O. Obie, C. Chua, I. Avazpour, M. Abdelrazek, J. Grundy, T. Bednarz, A framework for authoring logically ordered visual data stories, in: 2019 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC), 2019, pp. 257–259. `doi:10.1109/VLHCC.2019.8818925`.

[43] R. Liu, A low complexity topological sorting algorithm for directed acyclic graph, International Journal of Machine Learning and Computing 4 (2014) 194–197. `doi:10.7763/IJMLC.2014.V4.411`.

[44] D. Ren, B. Lee, M. Brehmer, N. H. Riche, Reflecting on the evaluation of visualization authoring systems : Position paper, 2018 IEEE Evaluation and Beyond - Methodological Approaches for Visualization (BELIV) (2018) 86–92.

[45] A. Satyanarayan, J. Heer, Lyra: An interactive visualization design environment, Computer Graphics Forum 33 (3) (2014) 351–360. `doi:10.1111/cgf.12391`.

[46] Data, Vega datasets, [Online.] Available: https://github.com/vega/vega/tree/master/docs/data, accessed January, 2019.

[47] Climate, Bureau of meteorology, [Online.] Available: http://www.bom.gov.au/climate/outlooks/#/temperature/summary, accessed January, 2019.