Supporting Story Synthesis: Bridging the Gap between Visual Analytics and Storytelling

Siming Chen, Jie Li, Gennady Andrienko, Natalia Andrienko, Yun Wang, Phong H. Nguyen, and Cagatay Turkay

(2018)

Background
Research Problem
Related Works
Bridging the Gap
Proposed Solution
Thoughts

Background

Research Problem
Related Works
Bridging the Gap
Proposed Solution
Thoughts

Visual Analytics

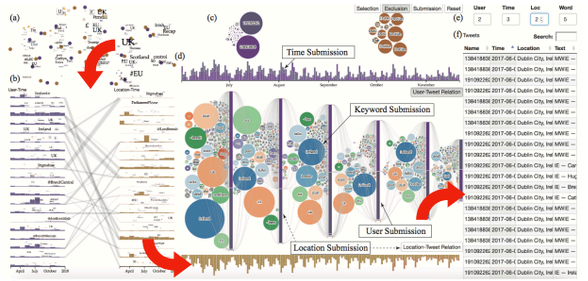
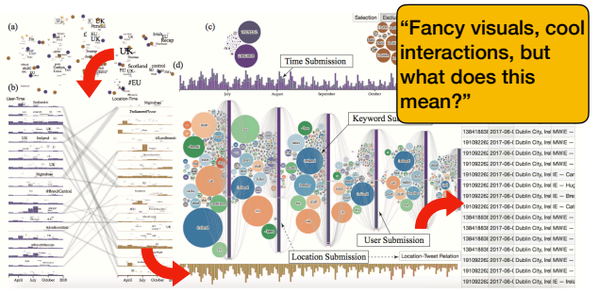


Figure 6. Visual analytics interface of social media, including (a) topic projection view, (b) topic comparison view, (c) user/ location distribution view, (d) temporal view, (e) story slice generation parameters, and (f) raw data table.

Visual Analytics



"Fancy visuals, cool interactions, but what does this mean?"

Figure 6. Visual analytics interface of social media, including (a) topic projection view, (b) topic comparison view, (c) user/ location distribution view, (d) temporal view, (e) story slice generation parameters, and (f) raw data table.

Storytelling

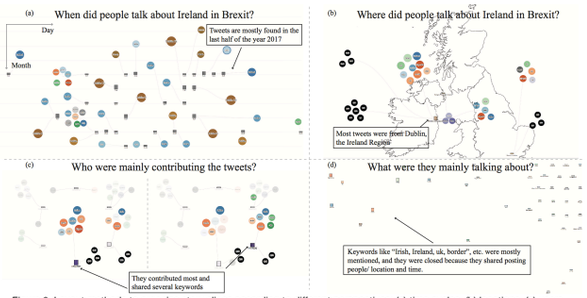
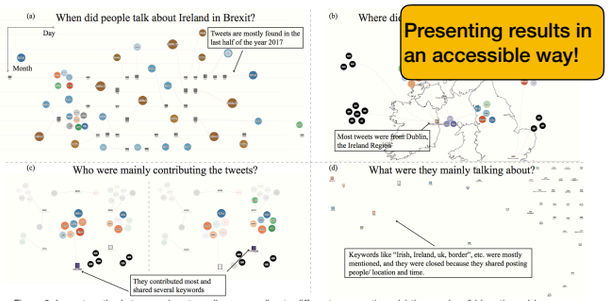


Figure 8. Layout methods to organize story slices according to different perspectives: (a) time cycles, (b) locations, (c) users, and (d) keywords.

Storytelling



Presenting results in an accessible way!

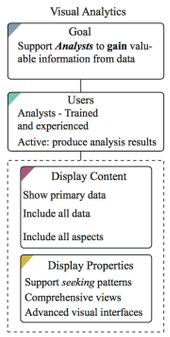
Figure 8. Layout methods to organize story slices according to different perspectives: (a) time cycles, (b) locations, (c) users, and (d) keywords.

Storytelling

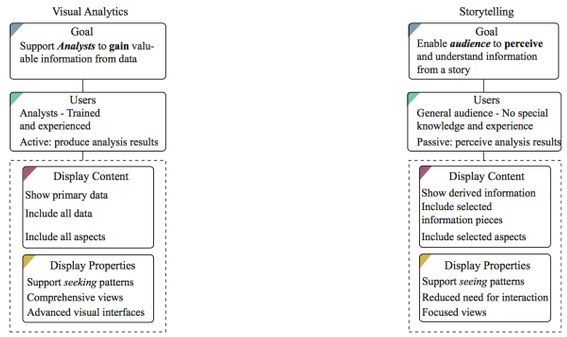
- Helps **non-experts** connect the dots
- Effectively **communicates insights**
- Provides **meaning** and **context** to **data**

Background
Research Problem
Research Problem
Related Works
Bridging the Gap
Proposed Solution
Thoughts

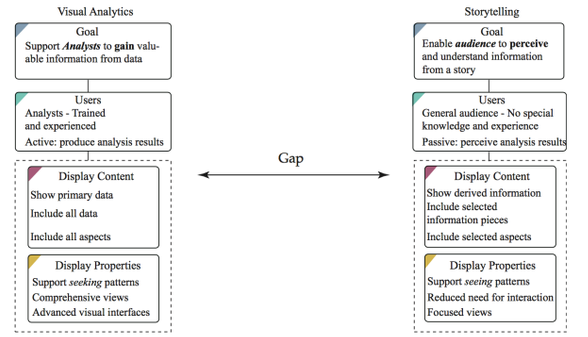
The Gap



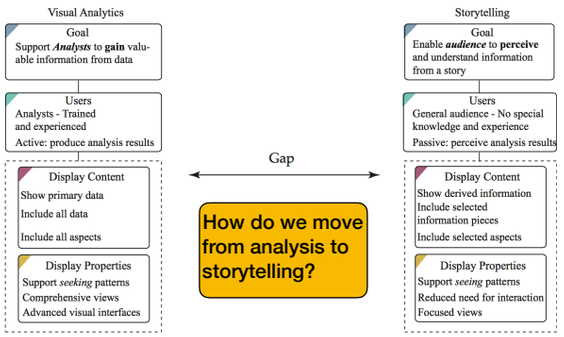
The Gap



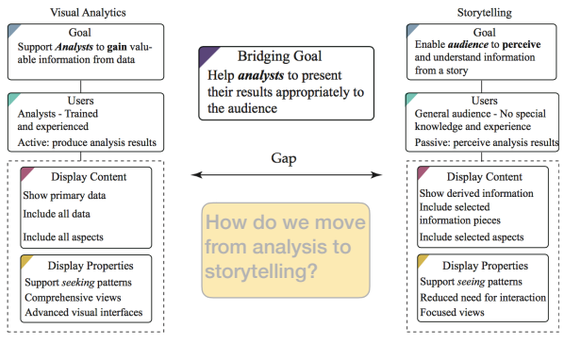
The Gap



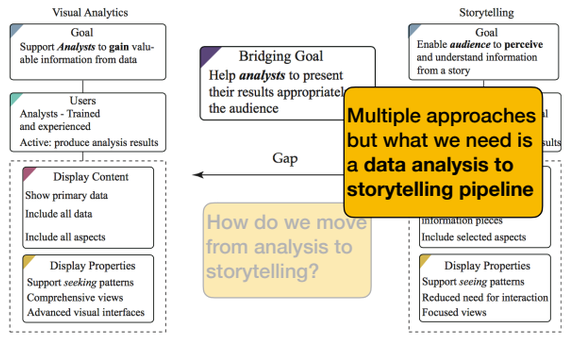
The Gap



The Gap



The Gap



Multiple approaches but what we need is a data analysis to storytelling pipeline

Background
Research Problem
Related Works
Bridging the Gap
Proposed Solution
Thoughts

Past Work

- Need for **organizing story contents**
- Existing works compose stories with **annotated screen shots** of **analytic displays**
- Provenance systems rely on **completed** analysis bookmarks
- Other systems rely on **completed** visualizations

Past Work

- Need for **organizing story contents**
- Existing works compose stories with **annotated screen shots** of **analytic displays**
- Provenance systems rely on **completed** analysis bookmarks
- Other systems rely on **completed** visualizations

Inherits visual complexity of analytic displays

Past Work

- Need for **organizing story contents**
- Existing works compose stories with **annotated screen shots** of **analytic displays**
- Provenance systems rely on **comple** bookmarks
- Other systems rely on **completed** visualizations

Inherits visual complexity of analytic displays

No current systems are integrated into the analysis workflow!

Background
Research Problem
Related Works
Bridging the Gap
Proposed Solution
Thoughts

Bridging the Gap

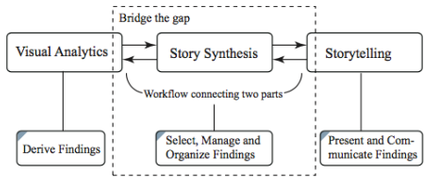


Figure 6. Proposed framework for bridging the gap.

Bridging the Gap

Extend analysis and storytelling workflows toward each other

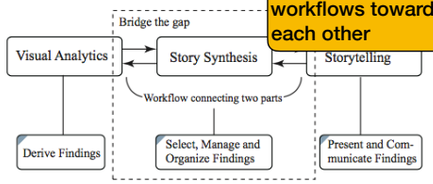


Figure 6. Proposed framework for bridging the gap.

Bridging the Gap

Story slice generation

Extend analysis and storytelling workflows toward each other

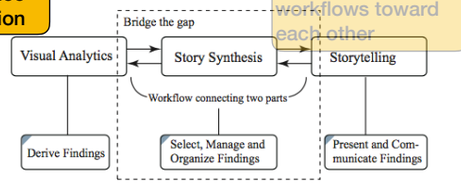


Figure 6. Proposed framework for bridging the gap.

Bridging the Gap

Story slice generation

Extend analysis and storytelling workflows toward each other

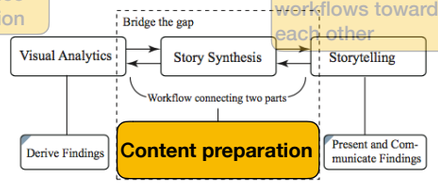


Figure 6. Proposed framework for bridging the gap.

Story Slices

- Structured representations of findings during **the course of analysis**
- Define what **facets** (information) **to capture**

VAST Challenge 2011 finding structure ::=

<label, time, location, N people, N messages, {(keyword, frequency)}, context>

Story Slices

- Structured representations of findings during **the course of analysis**
- Define what **facets** (information) **to capture**

VAST Challenge 2011 finding structure ::=

<label, time, location, N people, N messages, {(keyword, frequency)}, context>

VAST Challenge, analyze the circumstances of an epidemic outbreak via microblogs

Story Slices

- Structured representations of findings during **the course of analysis**
- Define what **facets** (information) **to capture**

VAST Challenge 2011 finding structure ::=

<label, time, location, N people, N me >

Epidemic and city were fictitious don't worry

VAST Challenge, analyze the circumstances of an epidemic outbreak via microblogs

Analytic Phase

Produce story slices during Analyze tasks

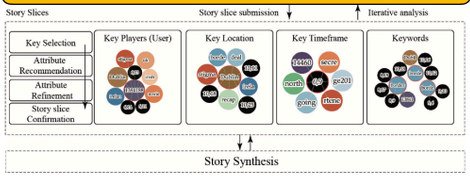


Figure 5. Visual analytics phase. After topic modelling, analysts explore the collection of messages from four perspectives and extract story slices.

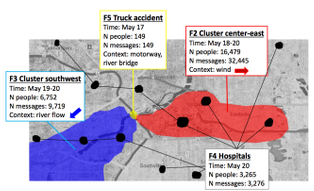
Story Slices

<F1 (outbreak), time = May 18-20, location = Vastopolis, N people = 27 446, N messages = 59 75
<F2 (cluster center-east), time = May 18-20, location = *polygon1*, N people = 16 479, N messages
<F3 (cluster southwest), time = May 19-20, location = *polygon2*, N people = 6 752, N messages =
<F4 (hospitals), time = May 20, location = {(-93.33, 42.24), (-93.42, 42.25), (-93.44, 42.20), ...}, N
<F5 (truck accident), time=May 17, location = (-93.427,42.226], N people = 149, N messages=149

...

...

69 755, {(chills, 10 436), (fever, 7 585), (sick, 6 543), ...}, null>
ages = 32 445, {(chills, 6 511), (fever, 4 905), ..., (flu, 3 466), ...}, context = {wind = west-to-east}>
es = 9 719, {(diarrhea, 2 785), (stomach, 2 682), ..., (nausea, 766), ...}, context = {river flow = north-to-southwest}>
}, N people = 3 265, N messages = 3 276, {{(chills, 1 419), (fever, 1 171), ..., (flu, 886), ...}, null>
=149, {(truck, 127), ..., (accident, 37), ..., (burning, 14), ..., (spilling cargo, 9), ...}, context = {motorway, river bridge}>



Story Slices

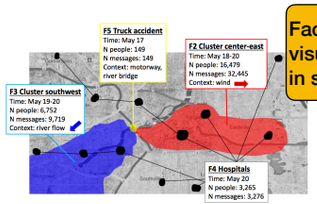
<F1 (outbreak), time = May 18-20, location = Vastopolis, N people = 27 446, N messages = 59 75
<F2 (cluster center-east), time = May 18-20, location = *polygon1*, N people = 16 479, N messages
<F3 (cluster southwest), time = May 19-20, location = *polygon2*, N people = 6 752, N messages =
<F4 (hospitals), time = May 20, location = {(-93.33, 42.24), (-93.42, 42.25), (-93.44, 42.20), ...}, N
<F5 (truck accident), time=May 17, location = (-93.427,42.226], N people = 149, N messages=149

...

...

69 755, {(chills, 10 436), (fever, 7 585), (sick, 6 543), ...}, null>
ages = 32 445, {(chills, 6 511), (fever, 4 905), ..., (flu, 3 466), ...}, context = {wind = west-to-east}>
es = 9 719, {(diarrhea, 2 785), (stomach, 2 682), ..., (nausea, 766), ...}, context = {river flow = north-to-southwest}>
}, N people = 3 265, N messages = 3 276, {{(chills, 1 419), (fever, 1 171), ..., (flu, 886), ...}, null>
=149, {(truck, 127), ..., (accident, 37), ..., (burning, 14), ..., (spilling cargo, 9), ...}, context = {motorway, river bridge}>

Facets help guide visual arrangements in story contents



Story Slices

<F1 (outbreak), time = May 18-20, location = Vastopolis, N people = 27 446, N messages = 59 75
<F2 (cluster center-east), time = May 18-20, location = *polygon1*, N people = 16 479, N messages
<F3 (cluster southwest), time = May 19-20, location = *polygon2*, N people = 6 752, N messages =
<F4 (hospitals), time = May 20, location = {(-93.33, 42.24), (-93.42, 42.25), (-93.44, 42.20), ...}, N
<F5 (truck accident), time=May 17, location = (-93.427,42.226], N people = 149, N messages=149

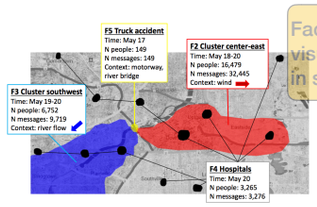
...

...

69 755, {(chills, 10 436), (fever, 7 585), (sick, 6 543), ...}, null>
ages = 32 445, {(chills, 6 511), (fever, 4 905), ..., (flu, 3 466), ...}, context = {wind = west-to-east}>
es = 9 719, {(diarrhea, 2 785), (stomach, 2 682), ..., (nausea, 766), ...}, context = {river flow = north-to-southwest}>
}, N people = 3 265, N messages = 3 276, {{(chills, 1 419), (fever, 1 171), ..., (flu, 886), ...}, null>
=149, {(truck, 127), ..., (accident, 37), ..., (burning, 14), ..., (spilling cargo, 9), ...}, context = {motorway, river bridge}>

Facets help guide visual arrang in story cont

Story slices are the input to the story synthesis phase



Facet Types

- **Discrete categories with no relationships**
- **Discrete entities with domain-specific relationships**
- **Linearly ordered elements**
- **Linearly ordered elements with distances**
- **Time**
- **1D/2D/3D space**

Facet Types

- **Discrete categories with no relationships**
- **Discrete entities with domain-specific relationships**
- **Linearly ordered elements**
- **Linearly ordered elements with distances**
- **Time**
- **1D/2D/3D space**

Facets specify possible arrangements and aggregations

Facet Types

- **Discrete categories with no relationships**
- **Discrete entities with domain-specific relationships**
- **Linearly ordered elements**
- **Linearly ordered elements with distances**
- **Time**
- **1D/2D/3D space**

These are very similar to the Abstract Data Types in VAD…

s specify possible gements and aggregations

Background
Research Problem
Related Works
Bridging the Gap
Proposed Solution
Thoughts

Story Synthesis

Story synthesis **creates story contents** from story slices by…

- **Aggregating** and **summarizing** story slices according to facets/measurements
- **Embedding** details into views/layouts
- **Arranging** story slices onto layouts according to facets/ relationships
- **Showing** facets by **priority** and from multiple perspectives
- **Annotating**

Story Synthesis

Story synthesis **creates story contents** from story slices by…

- **Aggregating** and **summarizing** story slices according to facets/measurements
- **Embedding** details into views/layouts
- **Arranging** story slices onto layouts according to facets/ relationships
- **Showing** facets by **priority** and from multiple perspectives
- **Annotating**

Story synthesis does not include the design of the final story!

Story Synthesis

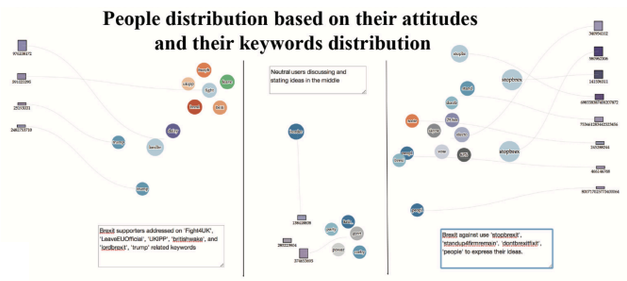


Figure 11. Output from story synthesis phase performed by sociology researcher.

Story Synthesis

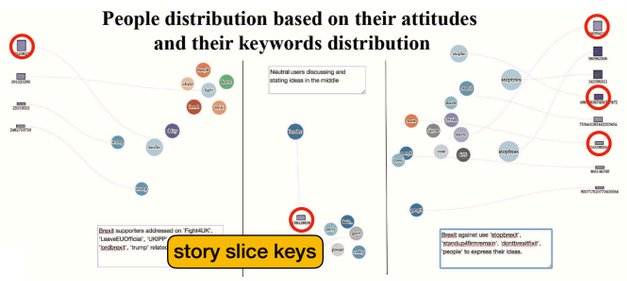


Figure 11. Output from story synthesis phase performed by sociology researcher.

Story Synthesis

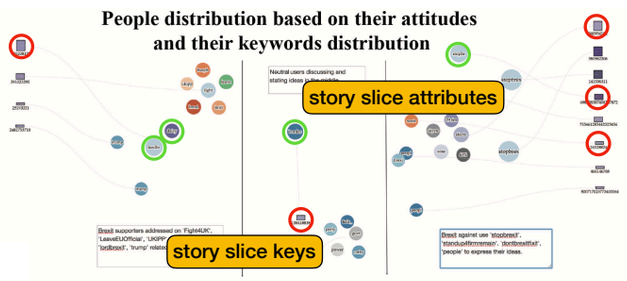


Figure 11. Output from story synthesis phase performed by sociology researcher.

Story Synthesis

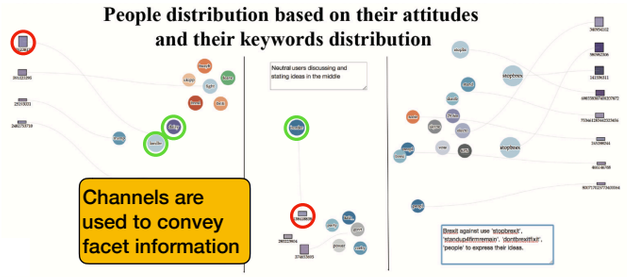


Figure 11. Output from story synthesis phase performed by sociology researcher.

New Framework

1. Define **types** and **structures** of story slices - What facets or patterns may be used from the data analysis phase?
2. Design **representations** of story slices - What data structure will be used to contain slice information? How will slices be visualized in the synthesis workspace?
3. Define story synthesis **support functions**.
4. Design the visual analytics system, including **support** for **discovery** of potential story slices.

New Framework

1. Define **types** and **structures** of story slices - What facets or patterns may be used from the data analysis phase?
2. Design **representations** of story slices - What data structure will be used to contain slice information? How will slices be visualized in the synthesis workspace?
3. Define story synthesis **support fun**
4. Design the visual analytics system, **discovery** of potential story slices.

This framework can be used in designing visual analytics systems that provide support for story synthesis.

Background
Research Problem
Related Works
Bridging the Gap
Proposed Solution
Thoughts

Good

☺ Need for framework is well backed by current research

☺ Story slices help reason about abstract data more

☺ Story slices capture model of analysis concept generation

☺ Evaluated by two sociology experts (positive feedback)

Bad

☹ Paper didn't do a great job proving generality of framework

☹ Majority of examples used online text message data

☹ Paper structure could be improved to help understanding

☹ Both experts only interviewed for 1.5h using similar data

☹ Is this framework better suited for messaging data?