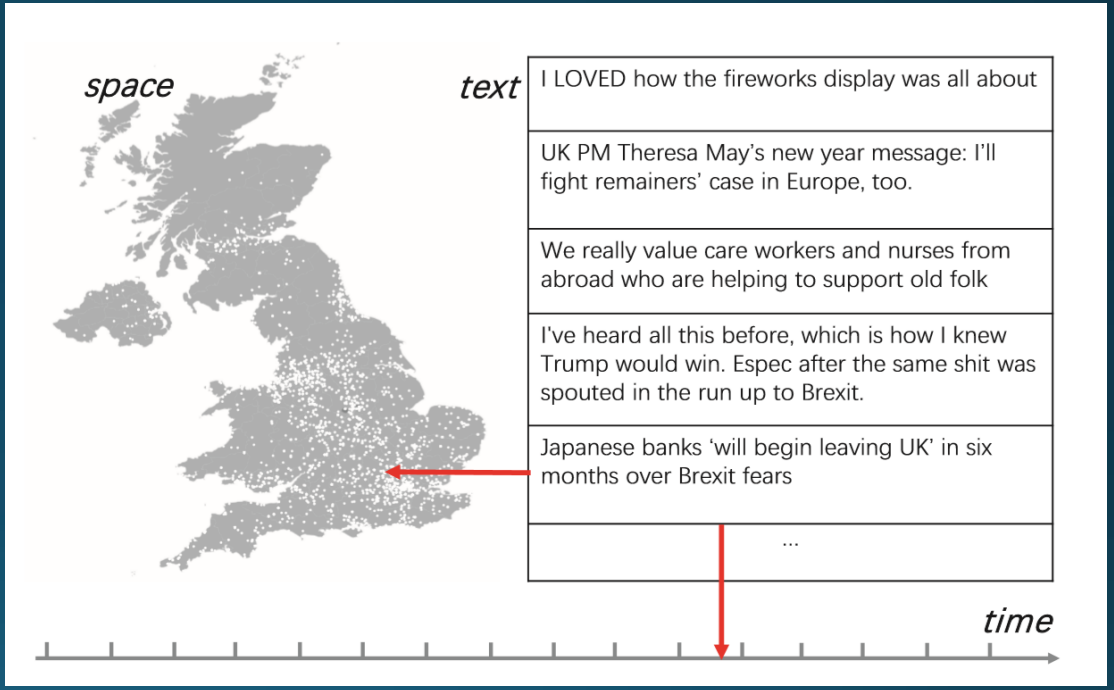
Semantics-Space-Time Cube: A Conceptual Framework for Systematic Analysis of Texts in Space and Time

Paper by Jie Li, Siming Chen, Wei Chen, Gennady Andrienko, and Natalia Andrienko.
IEEE Transactions on Visualization and Computer Graphics.

Presented by Mint Tanprasert
CPSC 547, Winter 2019, UBC

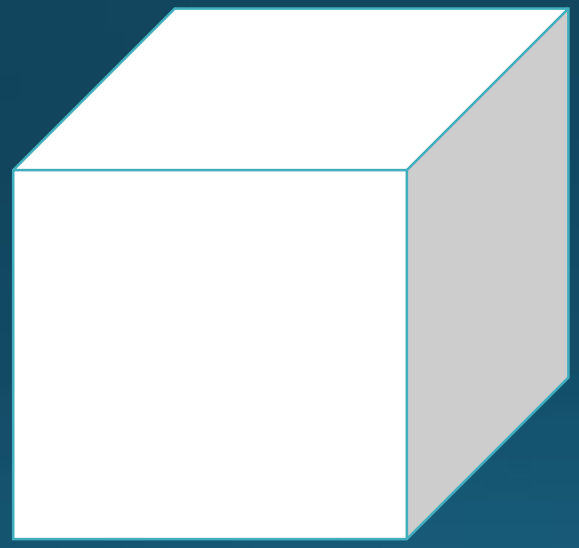
Text – Space – Time

- How does text semantics vary over space and time?
  - Explore complex relationship between the three facets
    - E.g. trend of topic popularity
  - Data example: geolocated social media posts



Proposed Visualization

"Cube"
metaphor

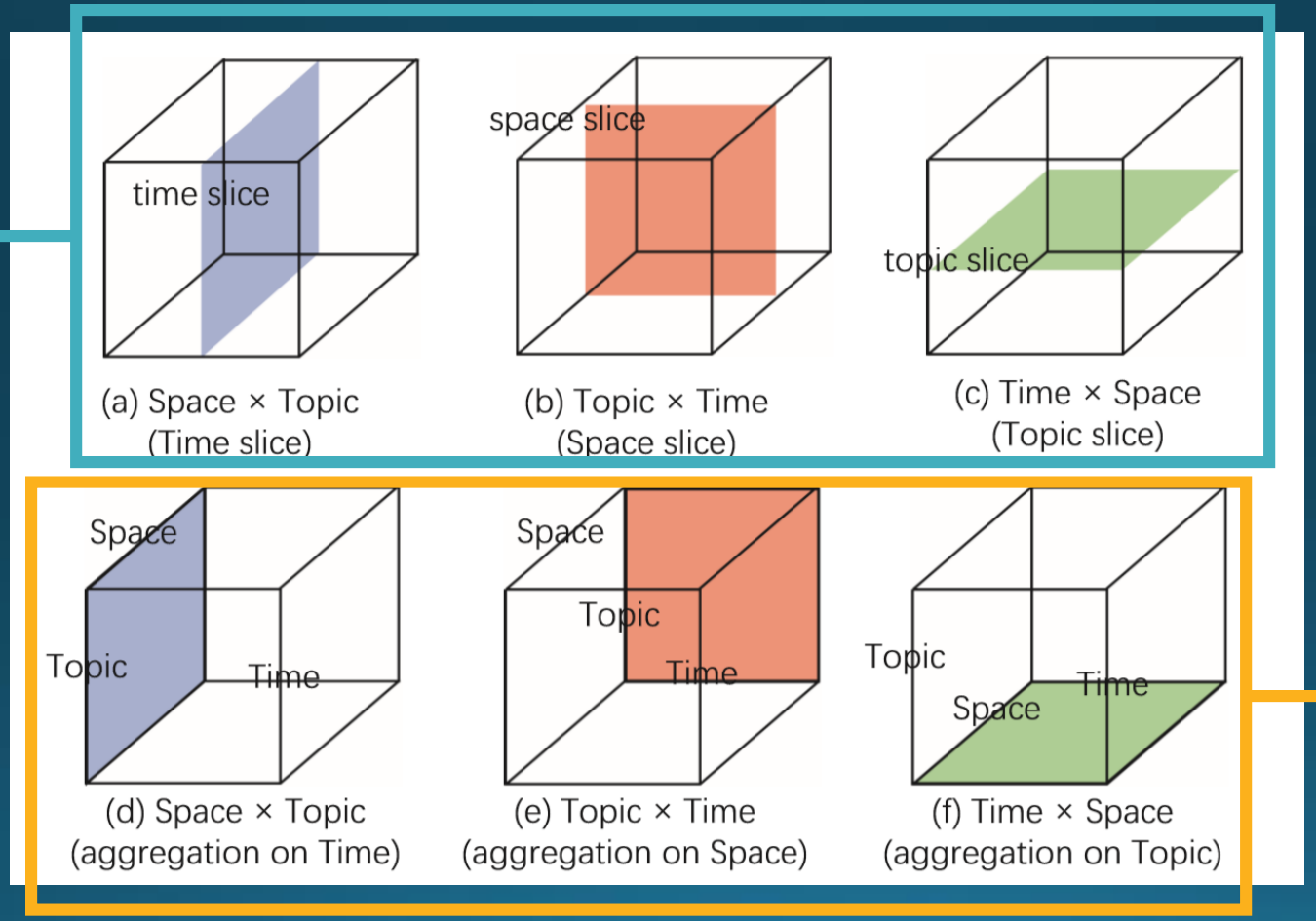


P x S x T

to**P**ics x **S**pace x **T**ime

Slices and Projections

Fix one variable

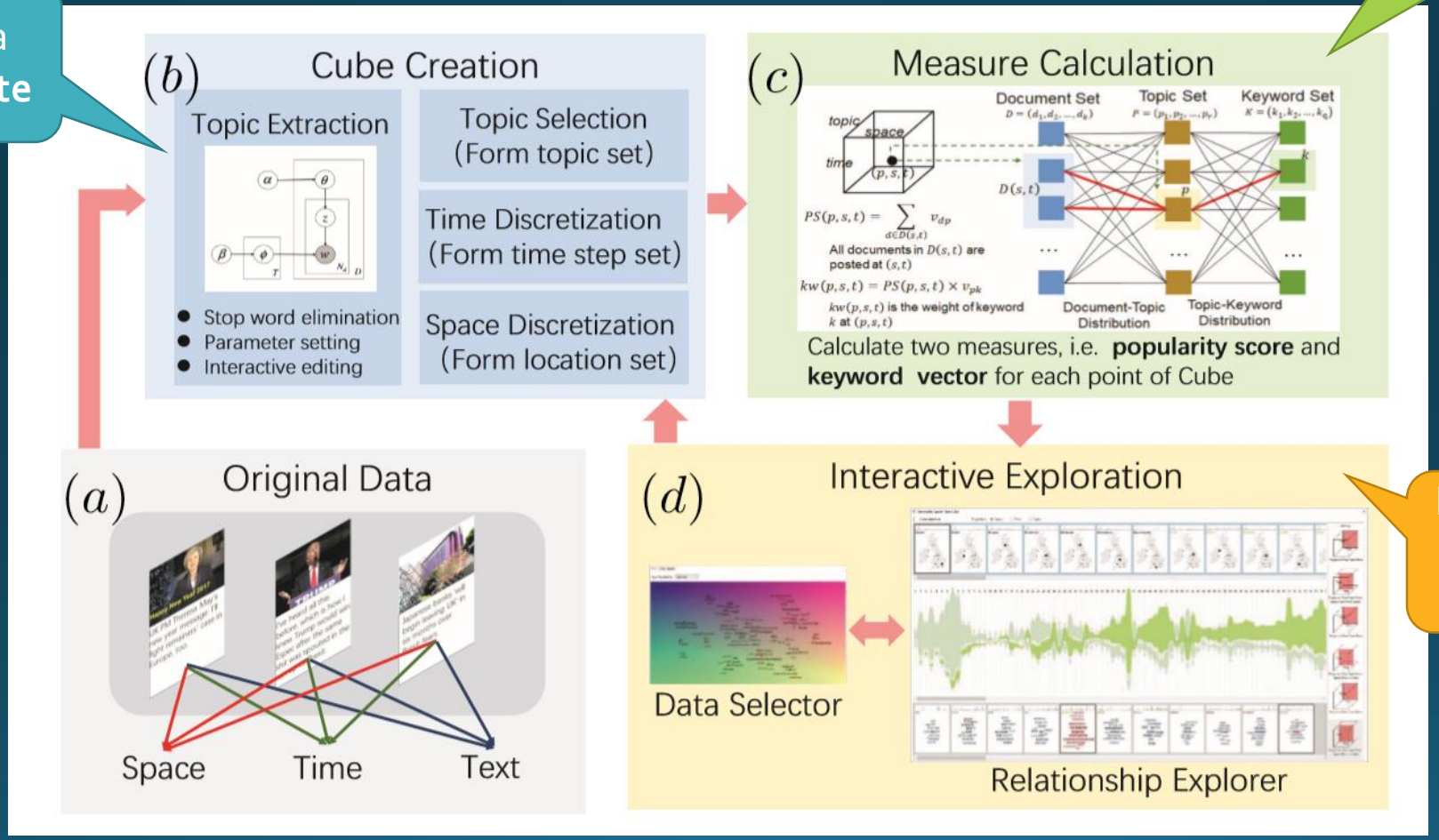


Aggregate over multiple slices (sum, mean, mode, quartiles, etc.)

Approach Overview

Make each data **discrete** and **finite**

Calculate **measures** for each point of the cube



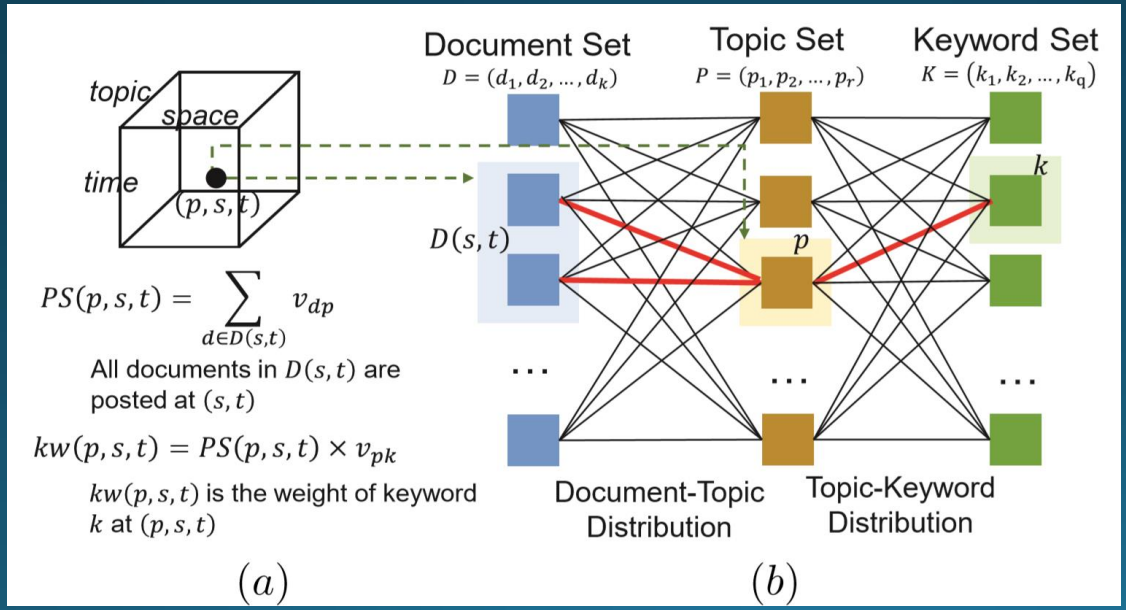
Design **interactive interface** for exploring data

Cube Creation

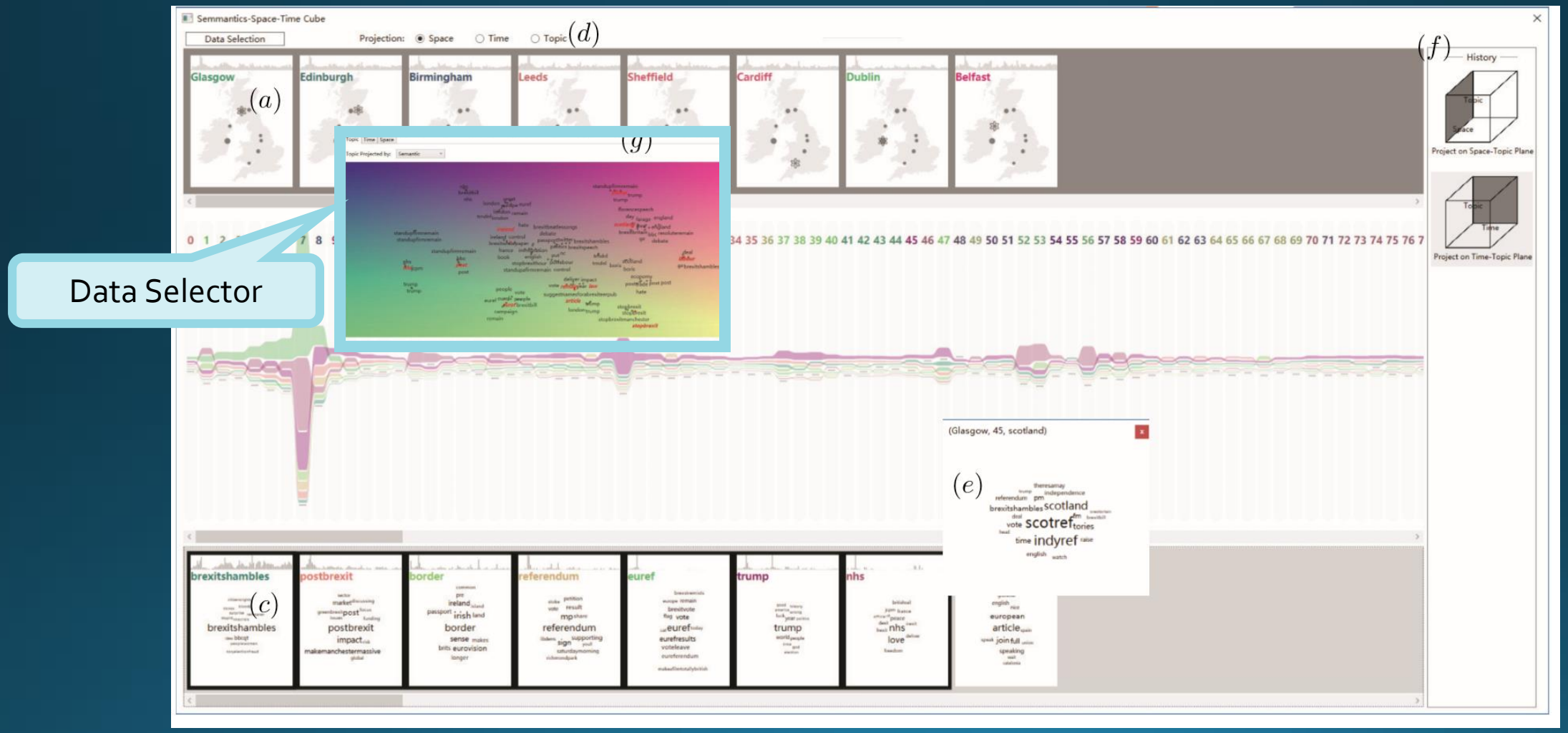
- **TEXT**: is represented by a vector of topic weights
  - *Topic* is defined as a probability distribution over a given set of *keywords*
  - Latent Dirichlet Allocation (LDA) – a probabilistic topic modeling method
  - Show all topics and let users select the ones they want to explore
- **TIME**: divide into intervals that are meaningful to humans e.g. weeks.
- **SPACE**: use individual and public activity locations, e.g. cities.

Measure Calculation

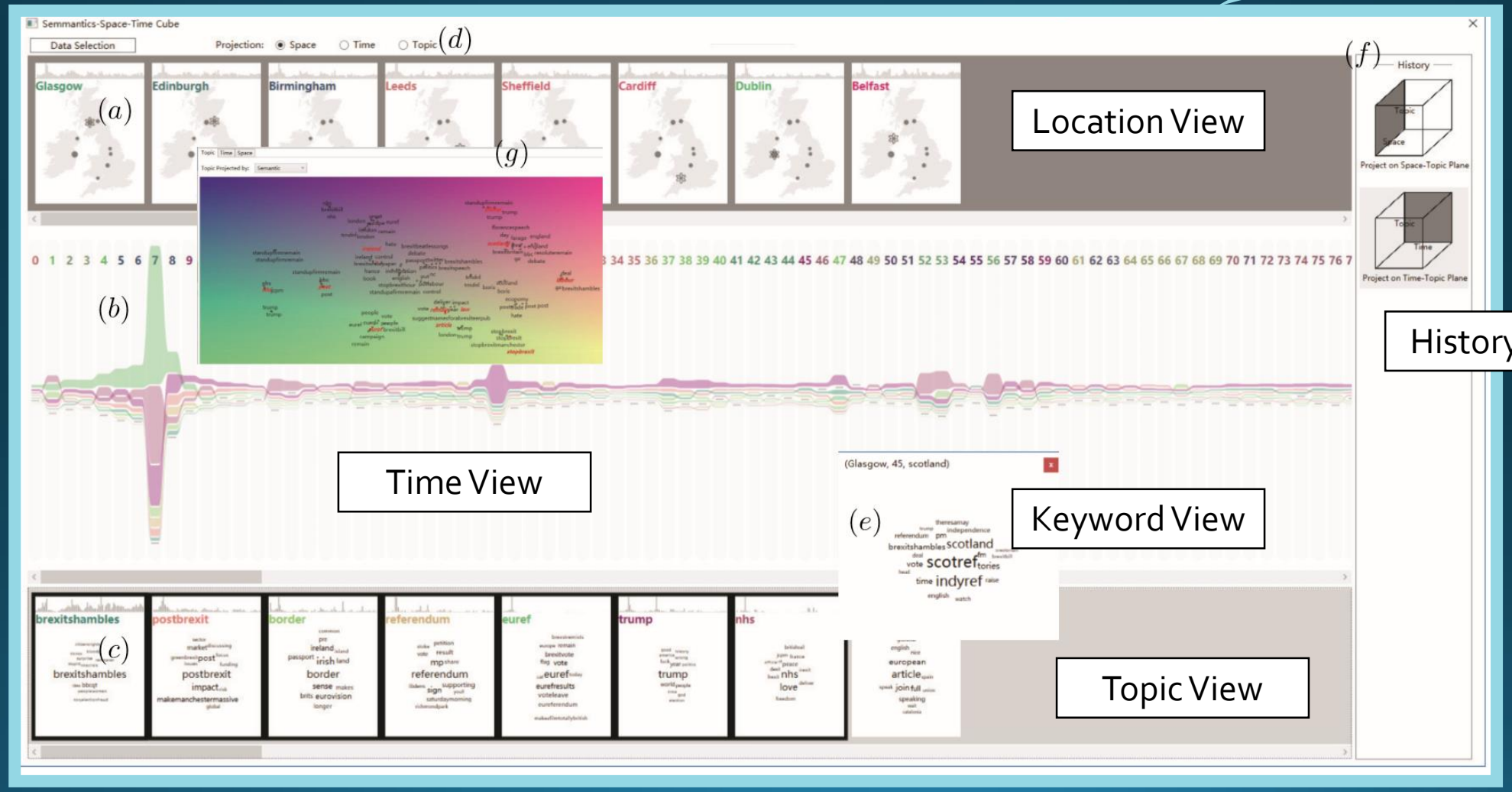
- At each *point (p, s, t)*, we can **derive** meaningful information
  - Popularity score
  - Keyword vector (list of <keyword, weight> pairs)



Interactive Exploration

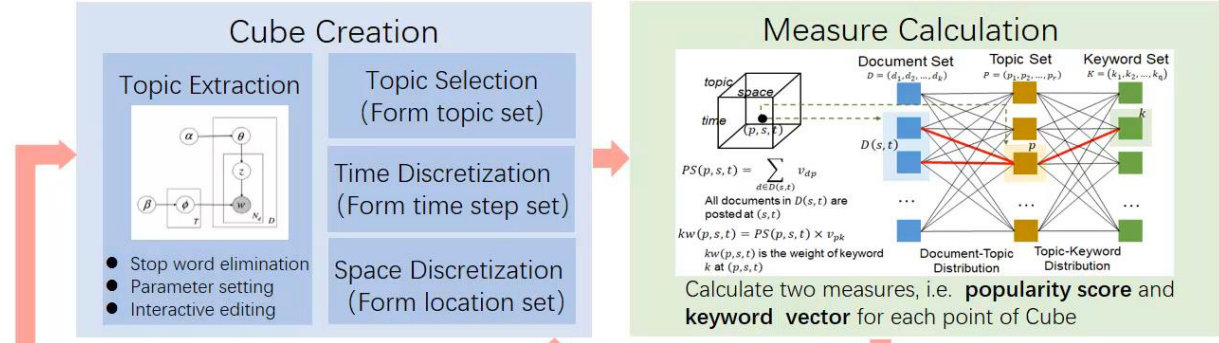


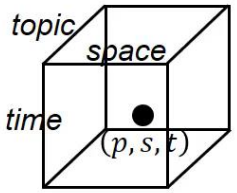
Interactive Exploration



Demo

Basic Idea





Goal: Exploration of various relationships among the space, time and topics.

T1: $t \rightarrow (S \rightarrow P \rightarrow M)$ or $\sum t \rightarrow (S \rightarrow P \rightarrow M)$

T2: $t \rightarrow (P \rightarrow S \rightarrow M)$ or $\sum t \rightarrow (P \rightarrow S \rightarrow M)$

T3: $s \rightarrow (T \rightarrow P \rightarrow M)$ or $\sum s \rightarrow (T \rightarrow P \rightarrow M)$

T4: $s \rightarrow (P \rightarrow T \rightarrow M)$ or $\sum s \rightarrow (P \rightarrow T \rightarrow M)$

T5: $p \rightarrow (S \rightarrow T \rightarrow M)$ or $\sum p \rightarrow (S \rightarrow T \rightarrow M)$

T6: $p \rightarrow (T \rightarrow S \rightarrow M)$ or $\sum p \rightarrow (T \rightarrow S \rightarrow M)$

https://ieeexplore.ieee.org/abstract/document/8540796/media#media

Analysis Summary: **WHAT**

- Data
  - A table where each item has 3 attributes
    - Text (arbitrary and unstructured)
    - Location (spatial)
    - Time (sequential)
- Derived
  - Discretized and finite representation of each attribute
  - Popularity score and keyword vector

Analysis Summary: **HOW**

- Encode
  - Color-code objects based on position in a projection
  - Use geographic map with glyphs
  - Size-code keyword based on importance
  - Indicate popularity score by linear ordering
- Manipulate
  - Pop-up window
  - 2D navigation
- Facet
  - Superimpose streams in time view (extended mode)
  - Linked navigation
- Reduce
  - Filter with “slices”
  - Aggregate with “projection”

Critiques – Strengths

- Good use of the metaphor
  - A good conceptual model for developing database operations (from paper)
- Filtering and aggregation allow for exploring large amount of data
- Responsive and well-organized view coordination
- Uniformity in visual design
- Flexible navigation that supports various topics and tasks
- Follow “Overview First, Zoom and Filter, Details on Demand”

Critiques – Weaknesses

- Cannot observe the variation for all 3 dimensions, simultaneously (must slice or project first)
- The "cube" dimensions can't be too big (from paper)
- Information loss in discretization (from paper)
- Limited number of "cards" can be shown at a time
- Automated item sorting could create change blindness
- Spatial positions of keywords in keyword view is meaningless
- "Number of specific mentioned keywords may not really reflect the public opinions." (Expert feedback from paper)

Thank you for your attention

Any questions?