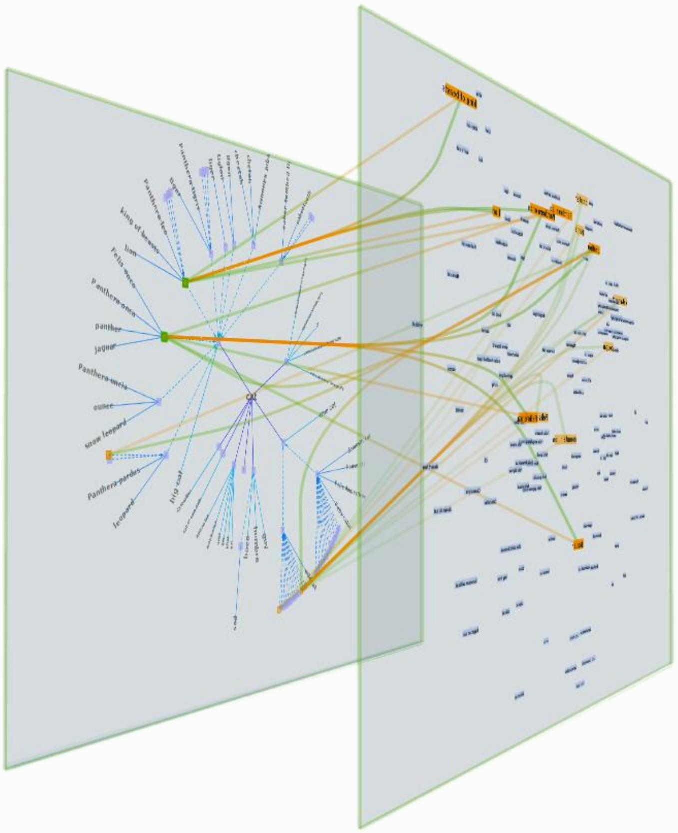
VisLink: Revealing Relationships Amongst Visualizations

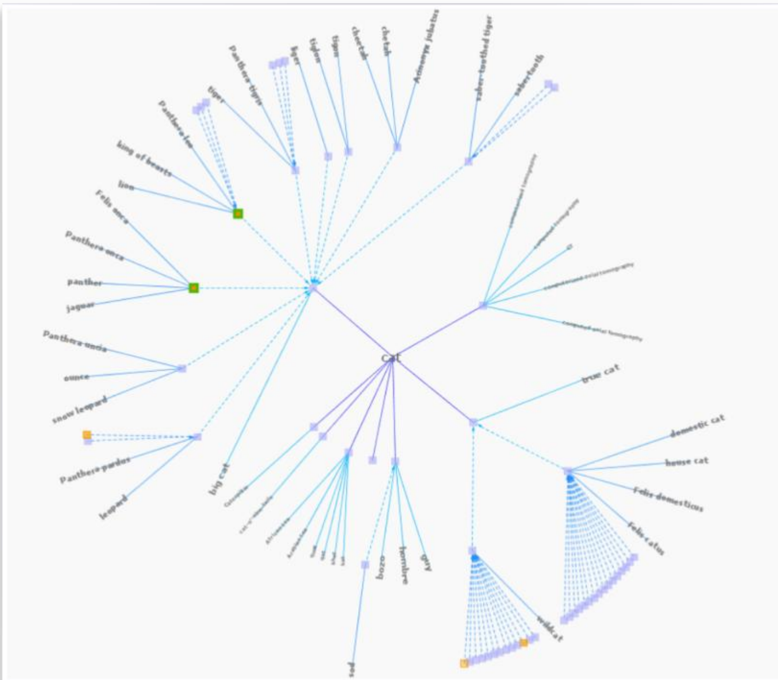
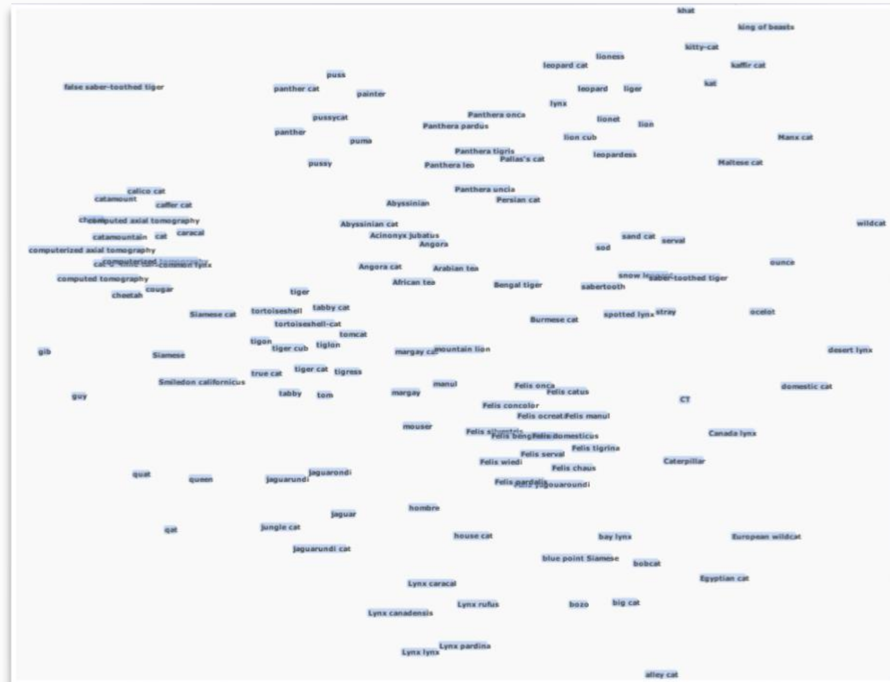
Christopher Collins
&
Sheelagh Carpendale

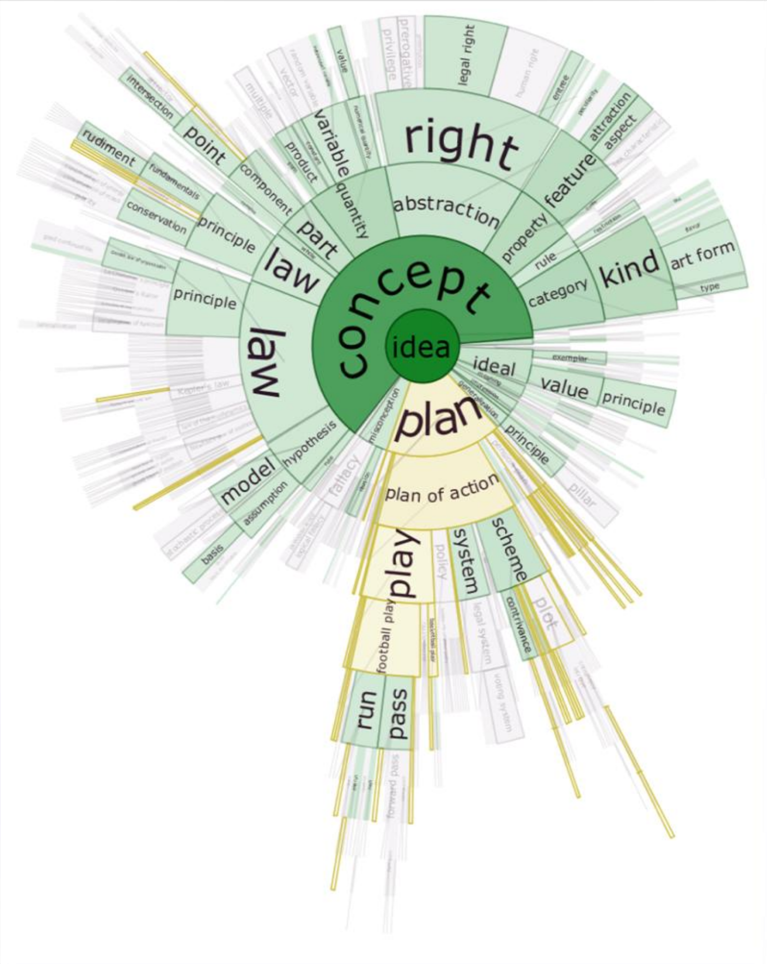
ccollins@cs.utoronto.ca, sheelagh@ucalgary.ca

University of Toronto

University of Calgary




cat
true cat
domestic cat
house cat
Felis domesticus
Felis catus
big cat
wildcat
king of beasts
lion
Panthera uncia
ounce
snow leopard
Panthera pardus
leopard
panther
jaguar
cheetah
saber-toothed tiger
sabertooth

false saber-toothed tiger
king of beasts
kitty-cat
kaffir cat
leopard cat
lioness
panther cat
painter
leopard
tiger
pussycat
Panthera onca
Panthera pardus
panther
puma
lion cub
Manx cat
Panthera tigris
Pallas's cat
leopardess
pussy
Panthera leo
Maltese cat
calico cat
catamount
caffer cat
Abyssinian
Panthera uncia
Persian cat
Abyssinian cat
Acinonyx jubatus
Angora
catamountain
caracal
computerized axial tomography
sand cat
serval
wildcat
Angora cat
Arabian tea
African tea
computed tomography
cougar
cheetah
Bengal tiger
sabertooth
ounce
tiger
Siamese cat
tabby cat
tortoiseshell-cat
spotted lynx
stray
ocelot
Burmese cat
tomcat
tigon
tiger cub
tiglon
mountain lion
desert lynx
Siamese
true cat
tiger cat
tigress
Smiledon californicus
manul
Felis onca
Felis catus
domestic cat
tabby
tom
margay
Felis concolor
CT
mouser
Felis domesticus
Canada lynx
Felis tigrina
Felis serval
Felis chaus
Caterpillar
queen
jaguarundi
Jaguarondi
jaguar
hombre
jungle cat
house cat
bay lynx
European wildcat
jaguarundi cat
blue point Siamese
bobcat
Egyptian cat
Lynx caracal
Lynx rufus
bozo
big cat
Lynx canadensis
Lynx lynx
Lynx pardina
alley cat



synsets, one for each sense of that word. Synsets in WordNet are related by many types of relationships, depending on the part of speech (noun, verb, etc.). WordNet contains 28 different types of relationships, but the most widely used part of WordNet is the hyponymy (`IS-A`) partial order. We remove all cycles (they are very rare) by taking a depth-first spanning tree at the user-selected root. In this work we focus on the noun hyponymy (`IS-A`) relationships in English WordNet (v2.1), rooted under the synset {*entity*} and having 73,736 nodes (synsets) and 75,110 edges. Verb hyponymy is also supported (but the tree is much smaller). The visualizations produced can be generalized to any partial order of a lexicon.

While development on WordNet continues, the interfaces for interacting with WordNet have not progressed to take advantage of advances in the field of information visualization. Currently available interfaces, both textual and graphical, focus on regions of local interest, for example by searching for the relationships for a single synset (ThinkMap, 2005; Bou, 2003; Alcock, 2004). In recent work, we created a working prototype of a visualization suite for WordNet which allows for an overview of the data, as well as the ability to focus on specific synsets of interest and obtain details. After developing these visualizations, we realized that the linguistic structure provided by WordNet could be useful not only for abstract visualization of the network itself, but that by applying other linguistic measures upon the nodes, we could better understand other aspects of language. Of particular interest to many in the information visualization and information retrieval communities is document structure and topic content.

In the following sections we will describe related work in document content visualization and present our interactive, animated, space-filling radial graph visualization of document content and WordNet hyponymy.

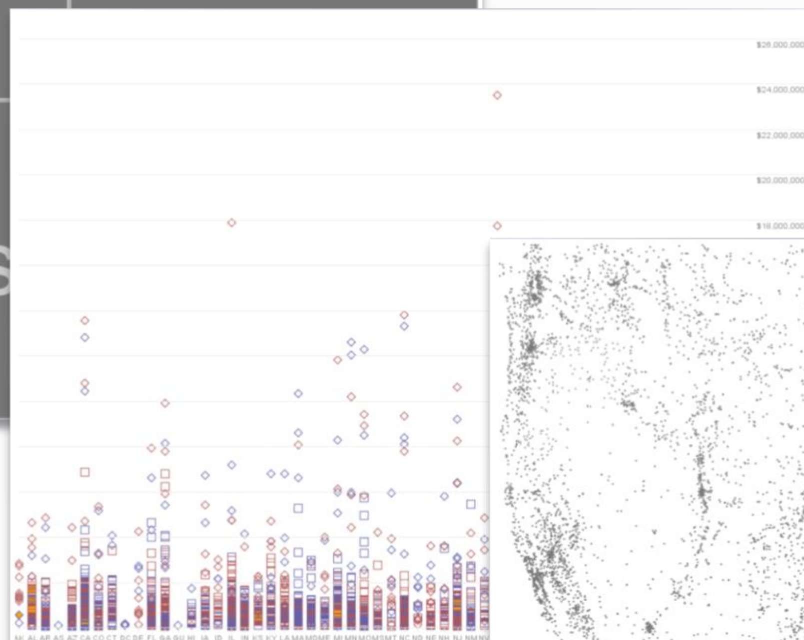
currence counts. Starstruck (Hetzler et al., 1998) creates glyphs by arranging lines in a circular pattern, where each line corresponds to a word and line length to word occurrence count. Gist Icons (DeCamp et al., 2005) builds on this idea using latent semantic indexing to group semantically-related search terms and reinforces the document as a glyph by drawing a smoothed iso-surface around the starstruck backbone. Blobby Text (Rohrer et al., 1999) is a 3D visualization which distorts a spherical surface to represent counts of up to 14 pre-selected terms of interest. Individual 3D glyphs are placed in the space according to their similarity. All three systems allow for inter-document comparison using arrays of glyphs, but do not investigate content within a single document.

Other visualizations of document content focus on the vocabulary and structure of a single document, such as TextArc (Paley, 2002), which arranges the sentences of a document in a circular layout with the individual words placed in the center. Self organizing (Kohonen) maps (Lin, 1992) have been used to reflect the relative strength of topics in a document, and the Document Lens (Robertson and Mackinlay, 1999) is an approach to focus-in-context distortion viewing of an entire document.

The document visualizations by Rembold and Späth (2006) compare and contrast themes and keywords within a collection of related documents while simultaneously revealing the thematic and typographic structure within an individual document. However, these visualizations are not interactive — they are printed graphics relating the essays in a collection. None of these approaches make use of formal linguistic structures such as WordNet.

TileBars (Hearst, 1995) is a document content visualization for information retrieval. It creates parallel small multiple visualizations, one for each user-specified search term set. Each small multiple consists of a bar divided into squares, each represent-

(Collins, 2007)

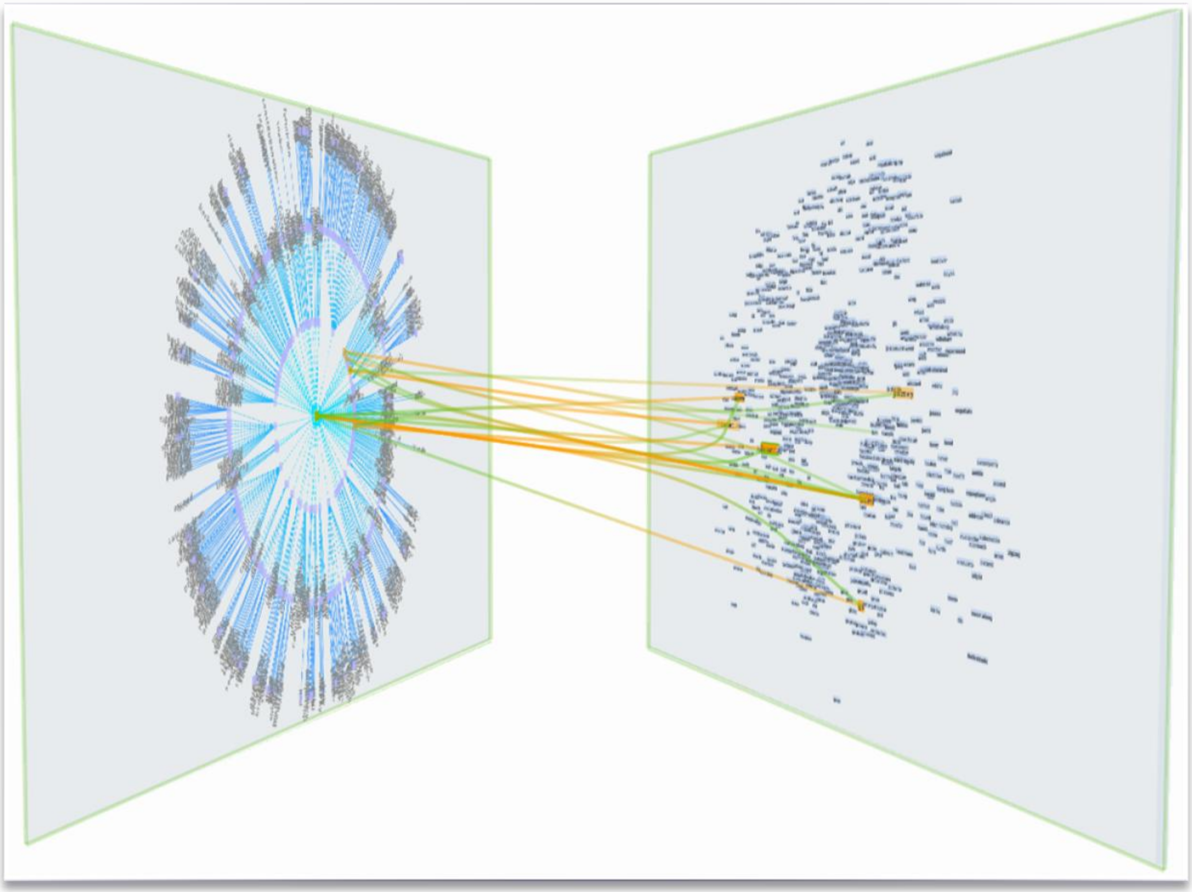


(Heer, 2006 [prefuse]) & (Fry, 2004)

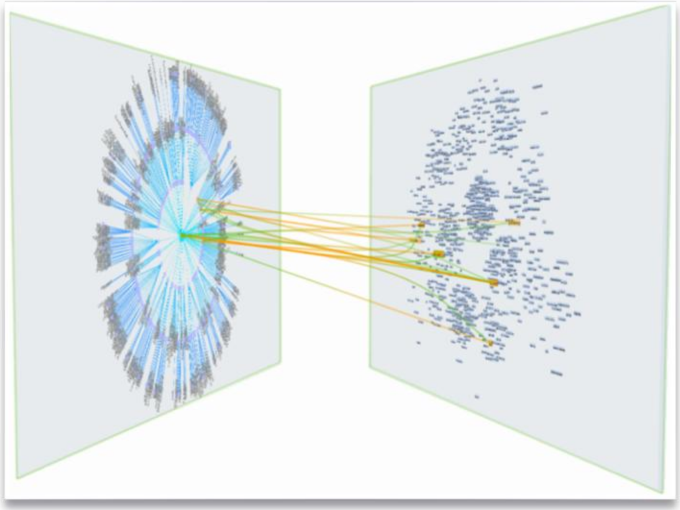
Understanding Multiple Relations

- What is the relationship…
  - across different views of the same data?
  - across different relations in the same dataset?
  - across multiple relations and datasets?

VisLink

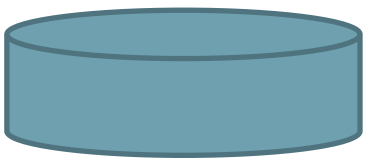


VisLink Overview

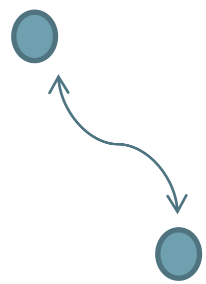


- Any number of 2D visualizations, each on its own plane in 3D space
- Adjacent planes connected by bundled edges
- Shortcuts and constrained widgets aid usability
- Enables powerful inter-visualization queries

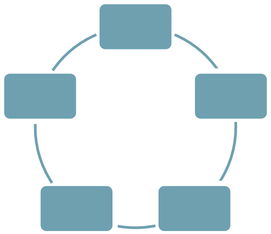
Formalizing Multiple Relations Visualizations



Dataset



Relation



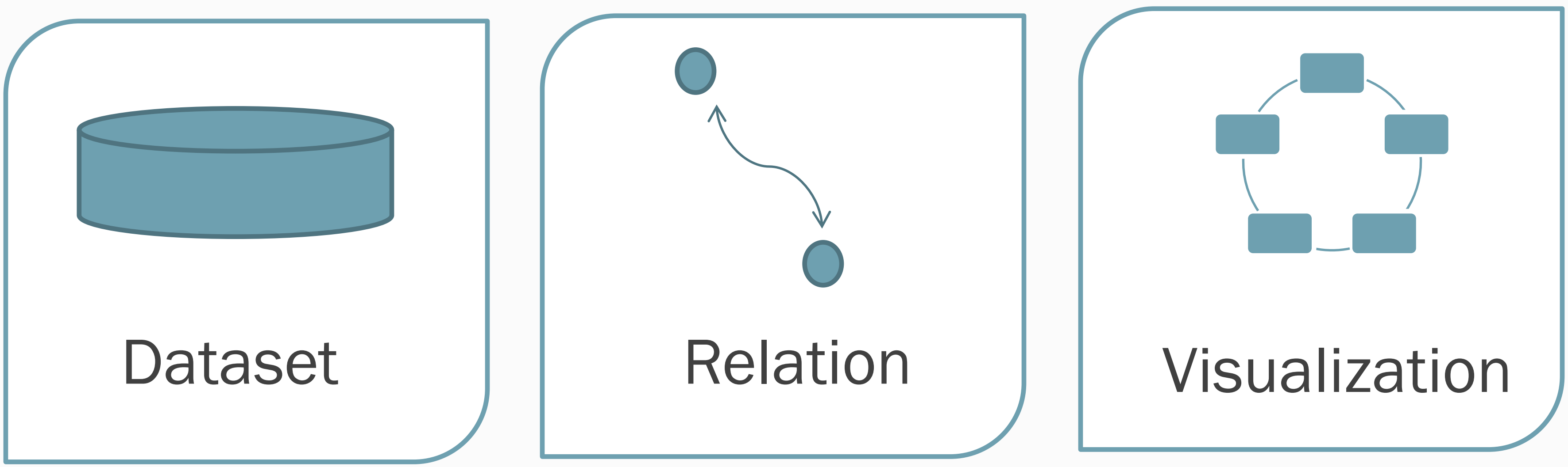
Visualization

Conference Attendee Data

Professor / Student

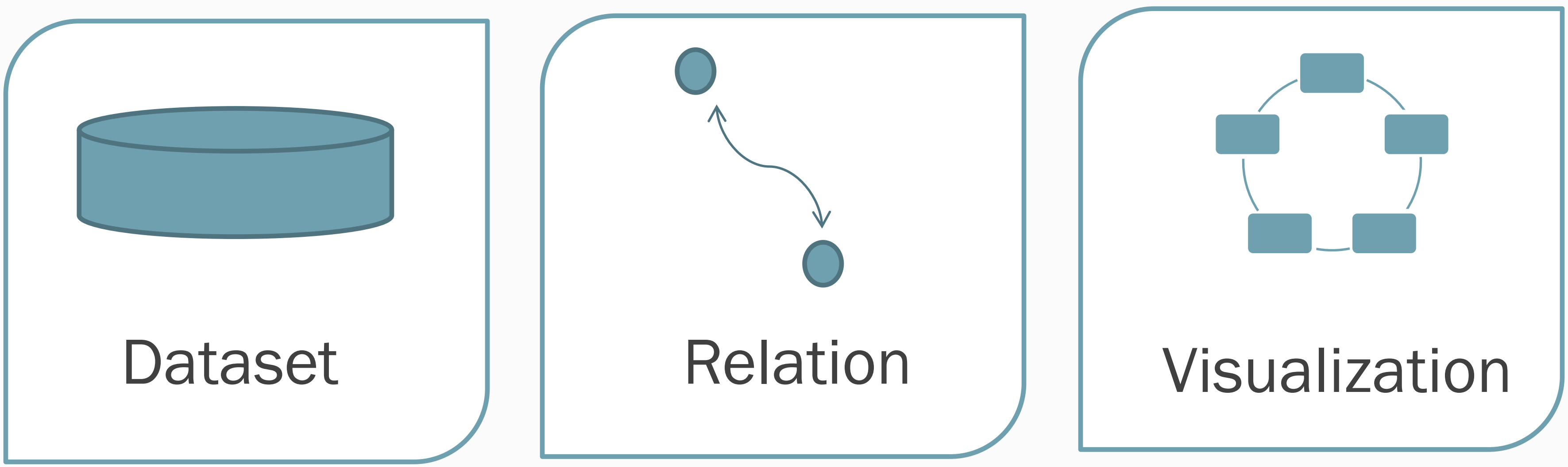
Node-link social network graph

Formalizing Multiple Relations Visualizations



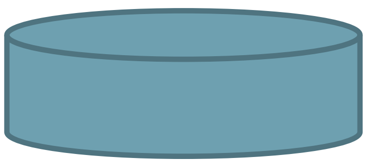
D_A

Formalizing Multiple Relations Visualizations



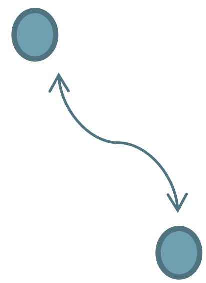
D_A $\quad$ $R_A(D_A)$

Formalizing Multiple Relations Visualizations



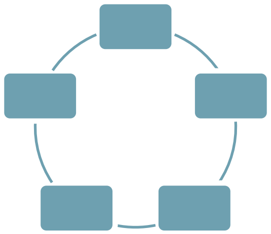
Dataset

D_A



Relation

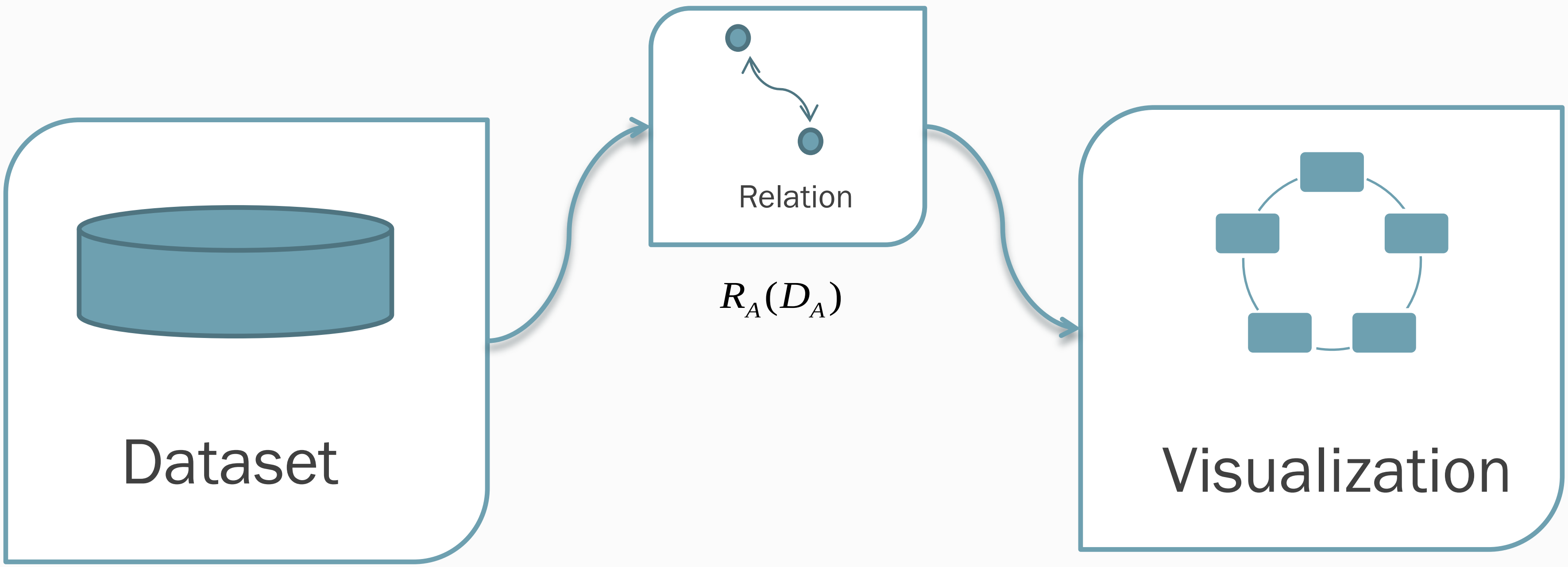
$R_A(D_A)$



Visualization

$Vis_A \rightarrow R_A(D_A)$

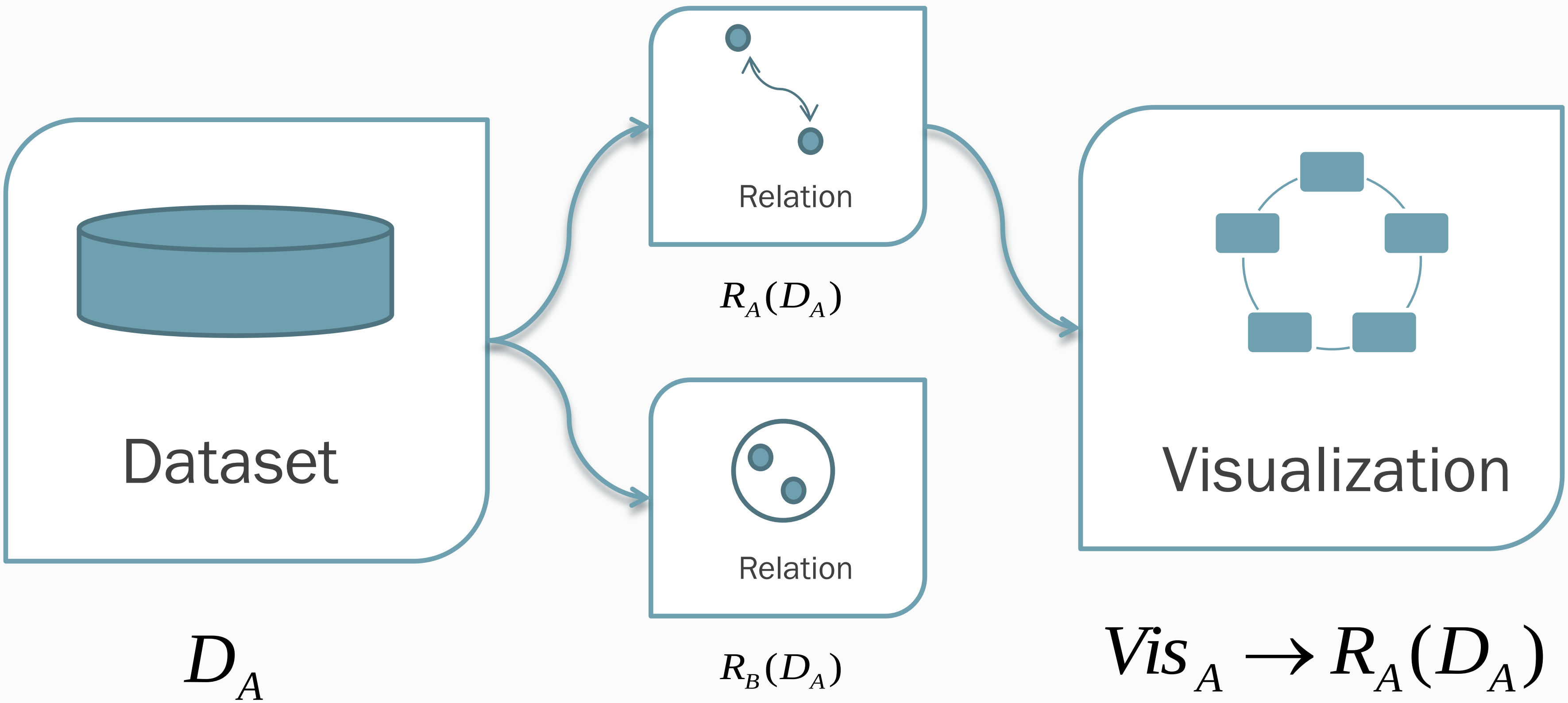
Formalizing Multiple Relations Visualizations



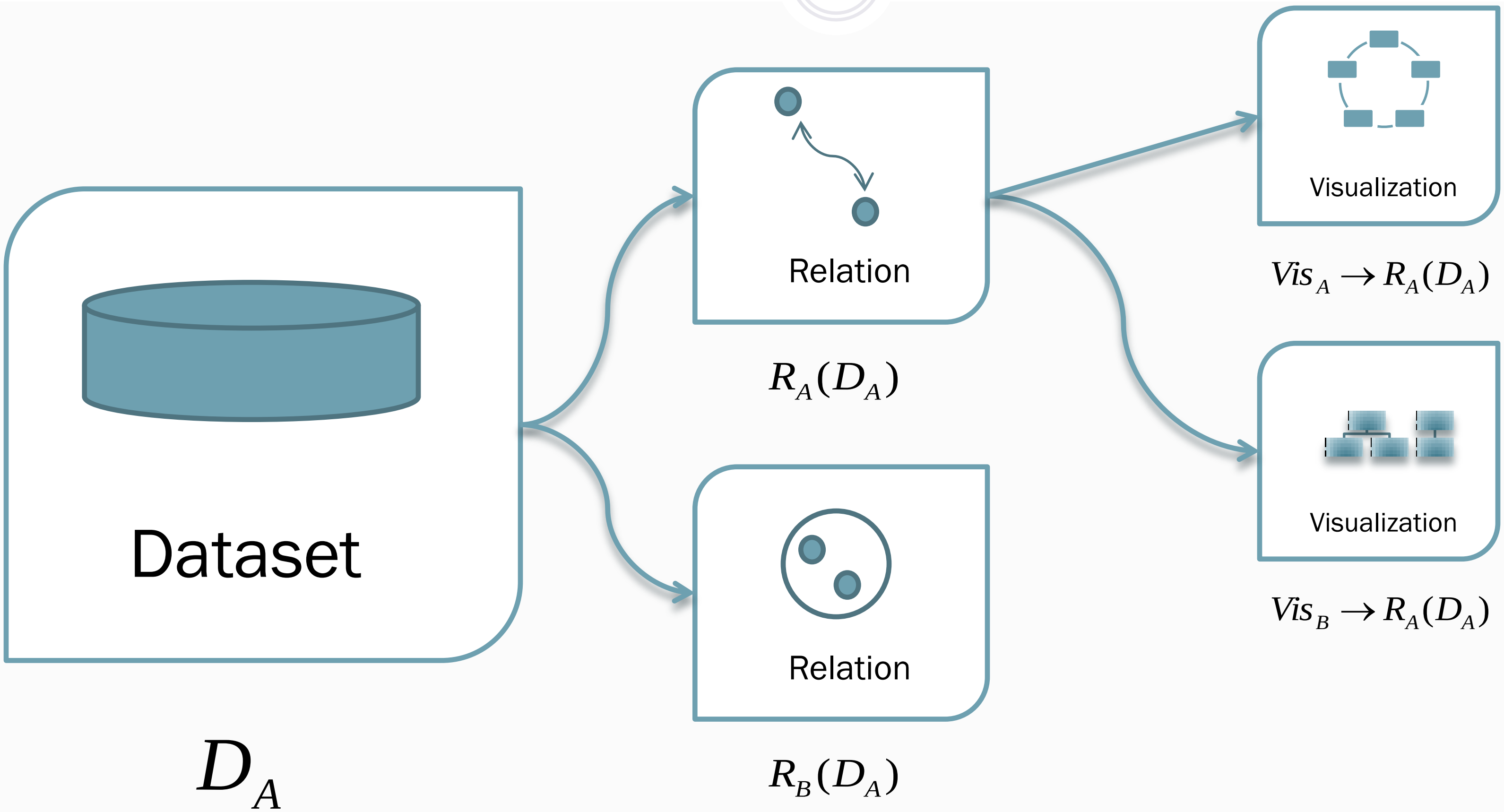
$$D_A$$

$$Vis_A \rightarrow R_A(D_A)$$

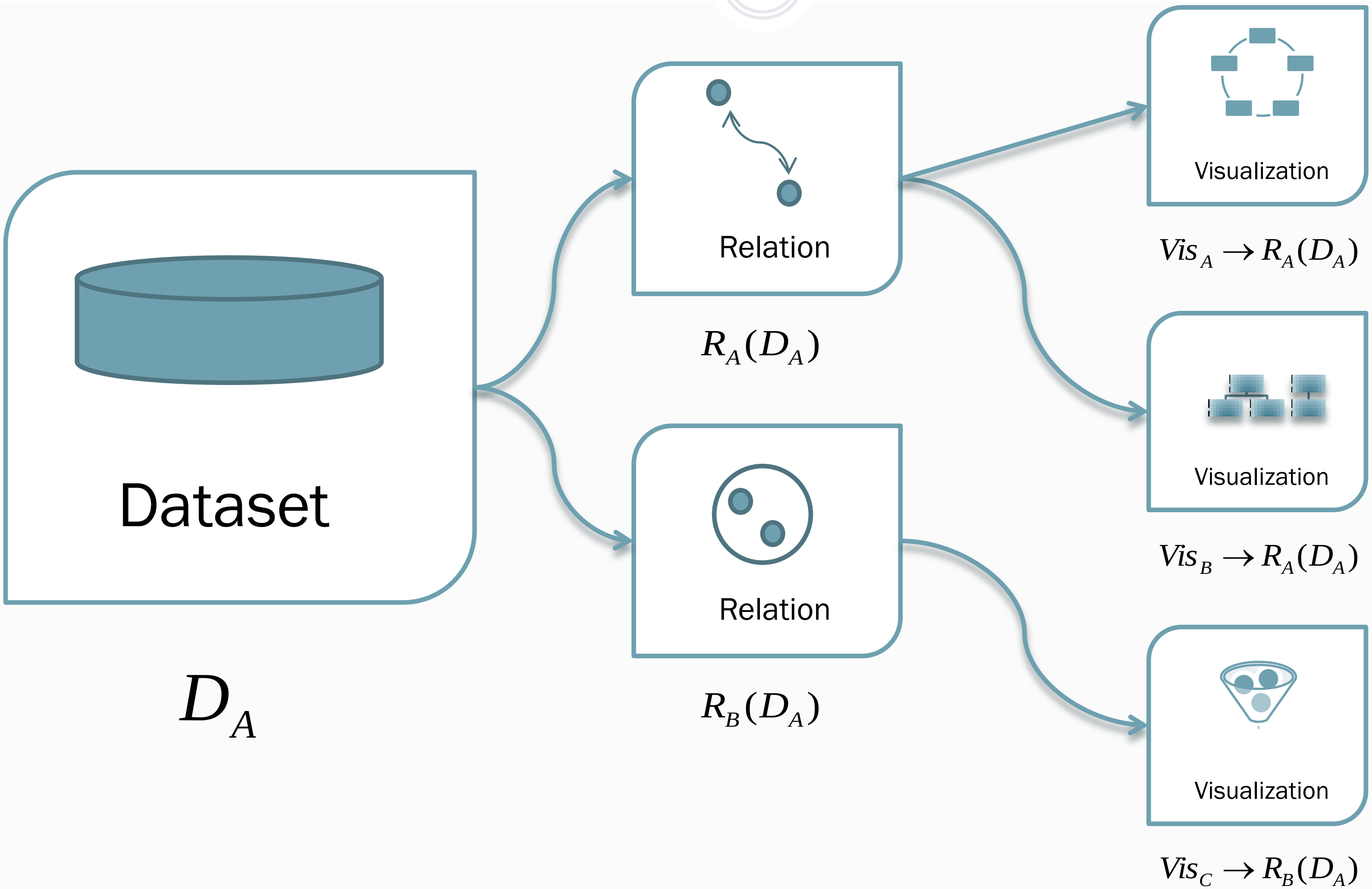
Formalizing Multiple Relations Visualizations



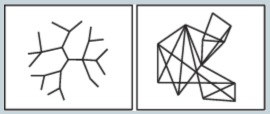
Formalizing Multiple Relations Visualizations



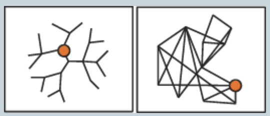
Formalizing Multiple Relations Visualizations



Multiple Relation Visualizations

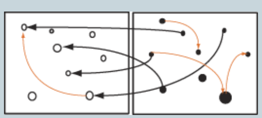


Individual Visualizations

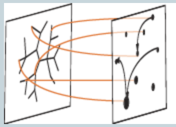


Coordinated Views

Compound Graphs

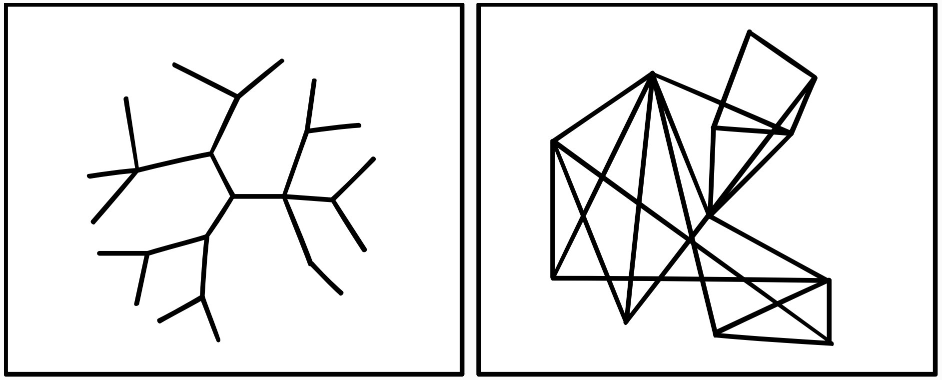


Semantic Substrates



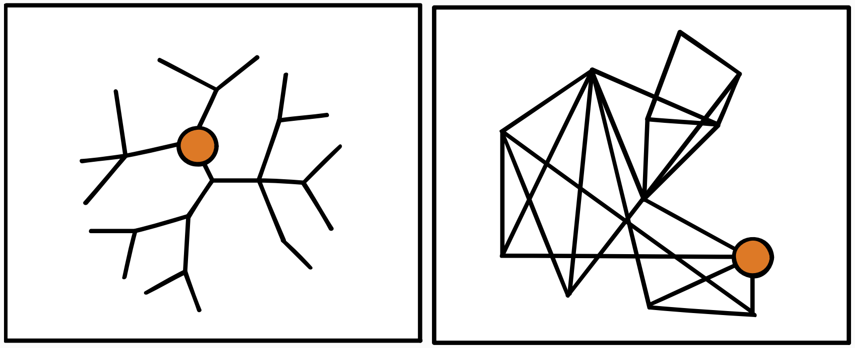
VisLink

Individual Visualizations

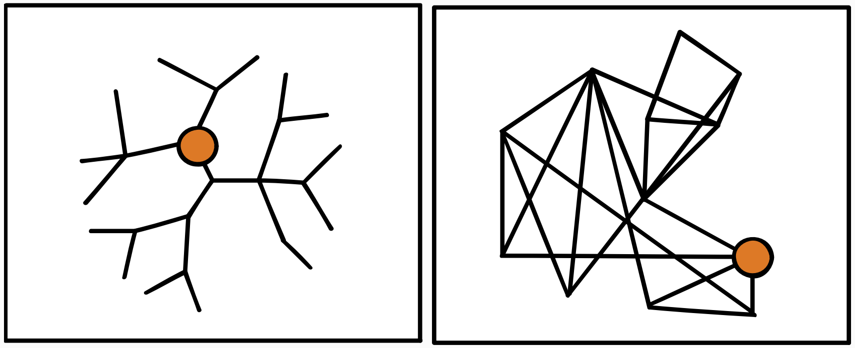


- Any datasets, relations, and visualizations
- Manually compare
- e.g. different charts in Excel

Coordinated Views

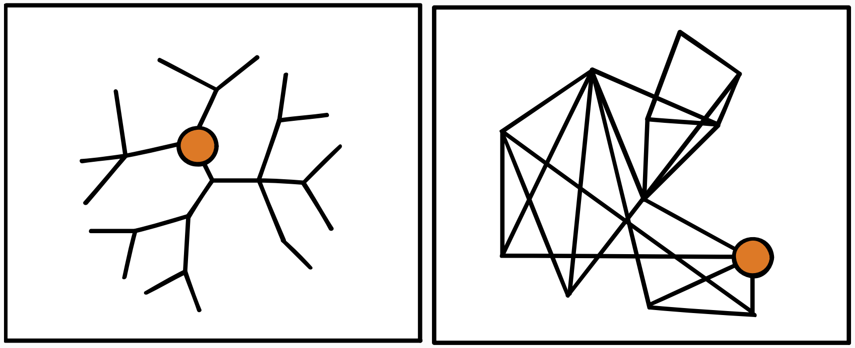


Coordinated Views



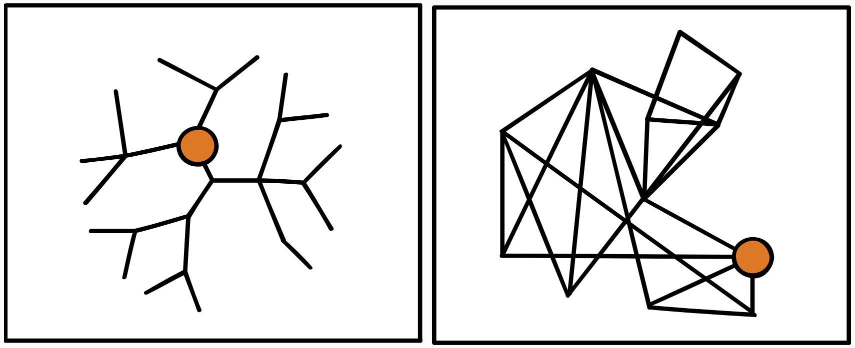
$$Vis_A \rightarrow R_A(D_A)$$

Coordinated Views



$$Vis_A \rightarrow R_A(D_A) \qquad\qquad Vis_A \rightarrow R_B(D_A)$$

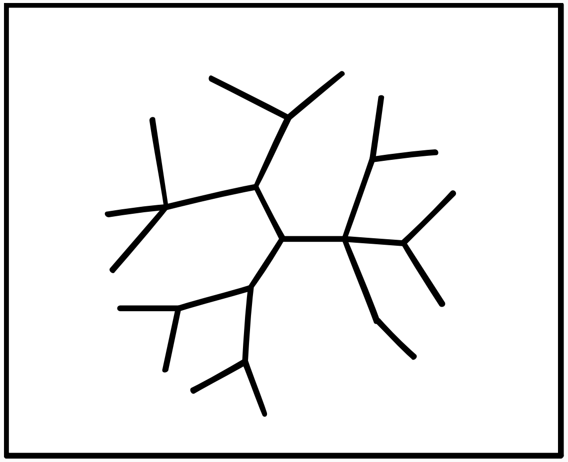
Coordinated Views



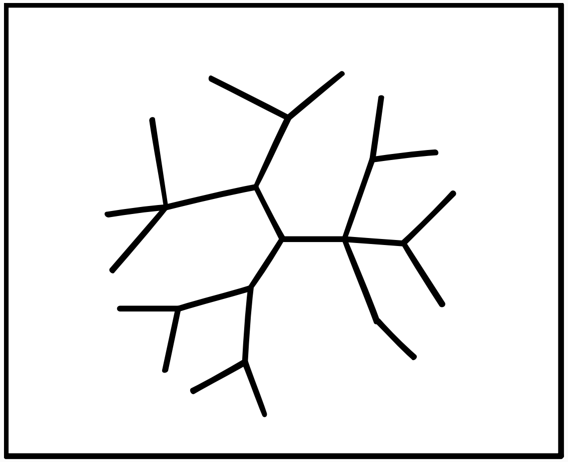
$$Vis_A \rightarrow R_A(D_A) \qquad Vis_A \rightarrow R_B(D_A)$$

- Any datasets, relations, and visualizations
- Interactive highlighting
- e.g., Snap-Together Visualization (North & Shneiderman, 2000)

Compound Graphs

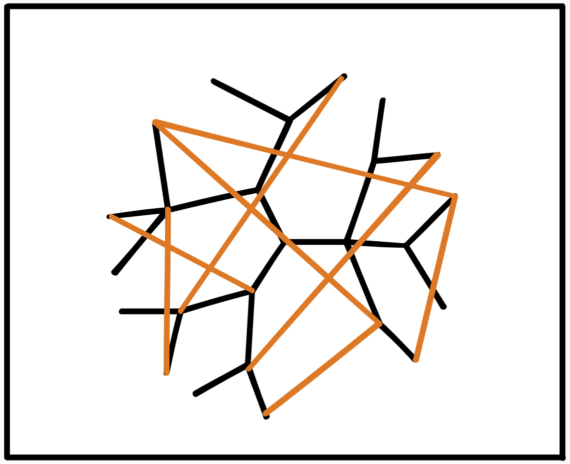


Compound Graphs



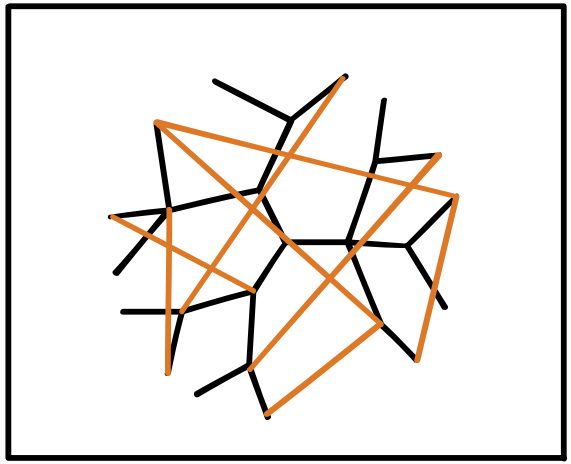
$$Vis_A \rightarrow R_A(D_A)$$

Compound Graphs



$$Vis_A \rightarrow R_A(D_A)\boxed{+R_B(D_A)}$$

Compound Graphs



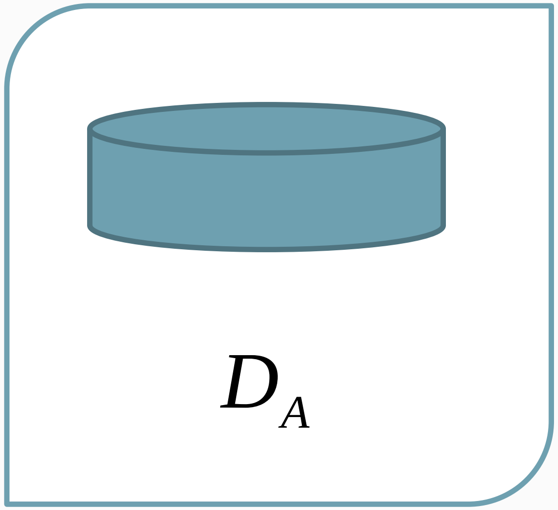
$$Vis_A \rightarrow R_A, R_B(D_A)$$

- Secondary relation has no spatial rights
- e.g., Overlays on Tr
  (Neumann et al., 2005),

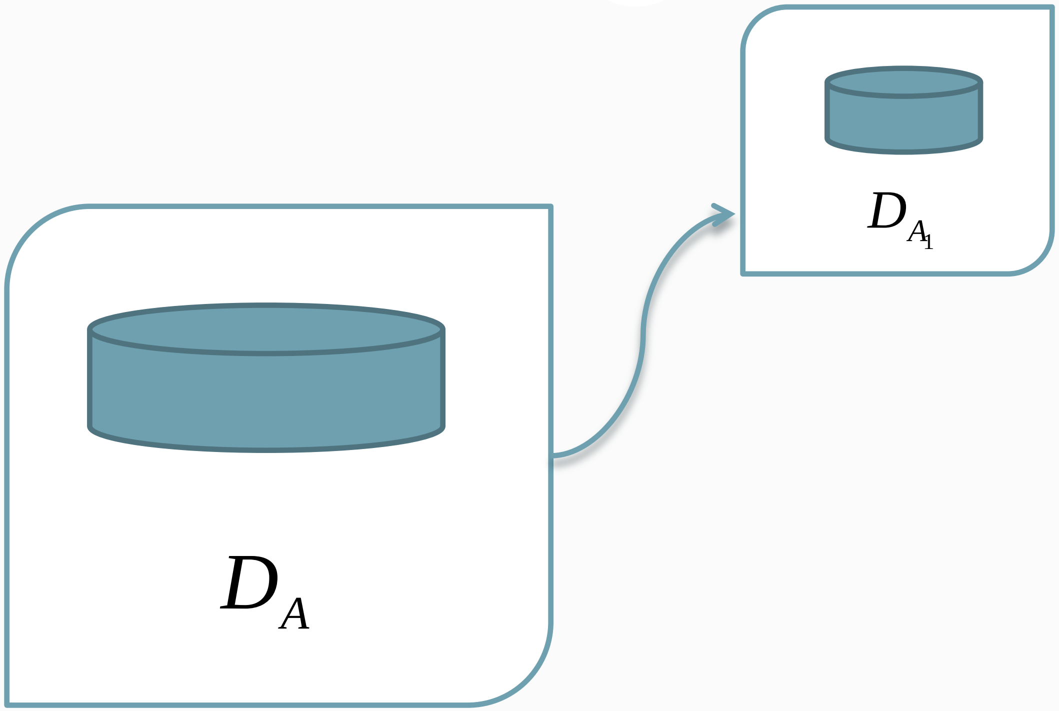
Use of the powerful spatial dimension to encode data relationships.

Semantic Substrates

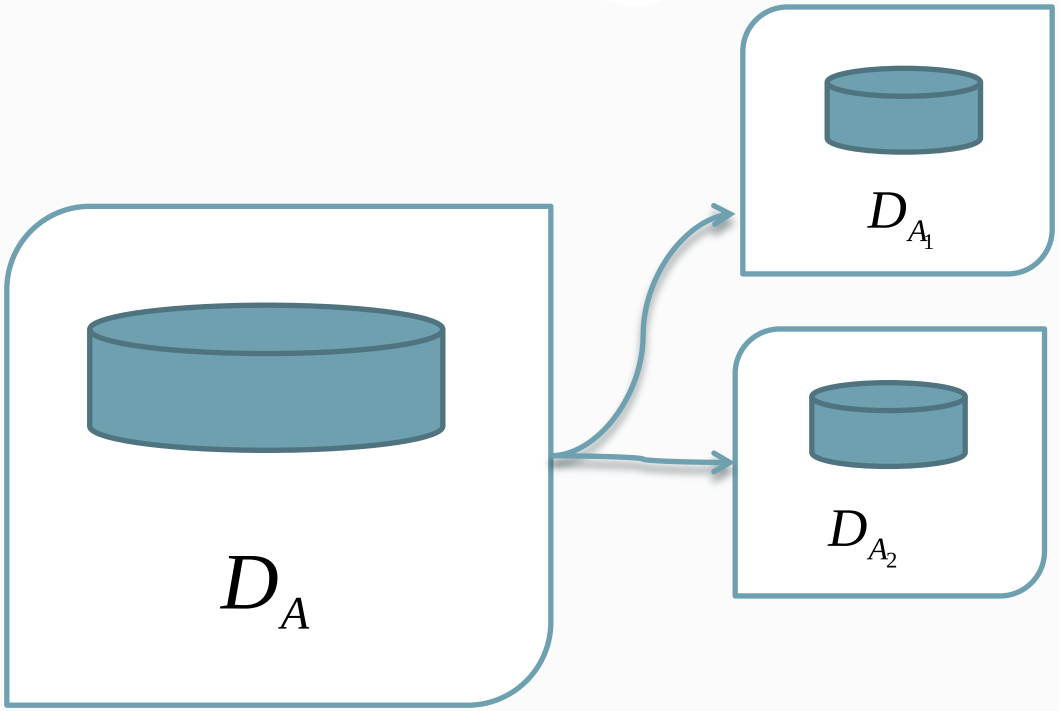
Semantic Substrates



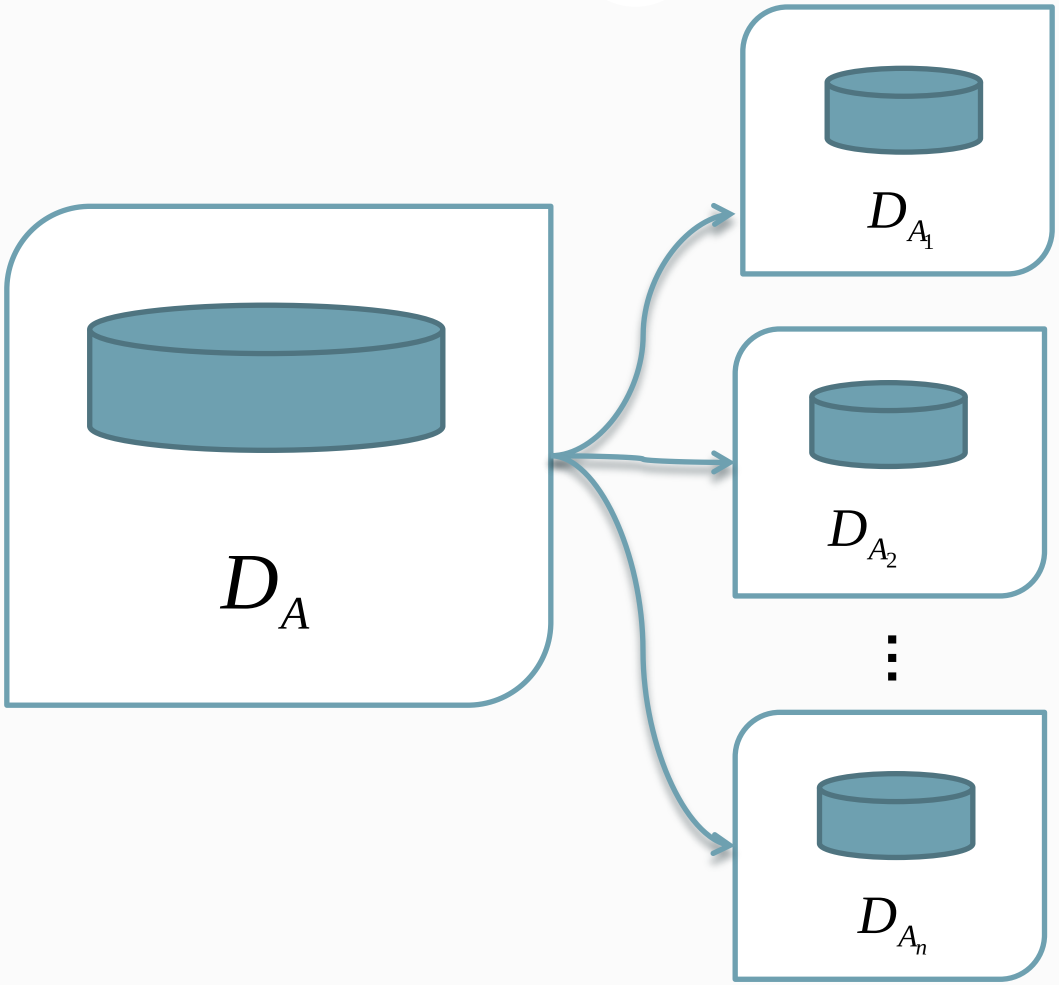
Semantic Substrates



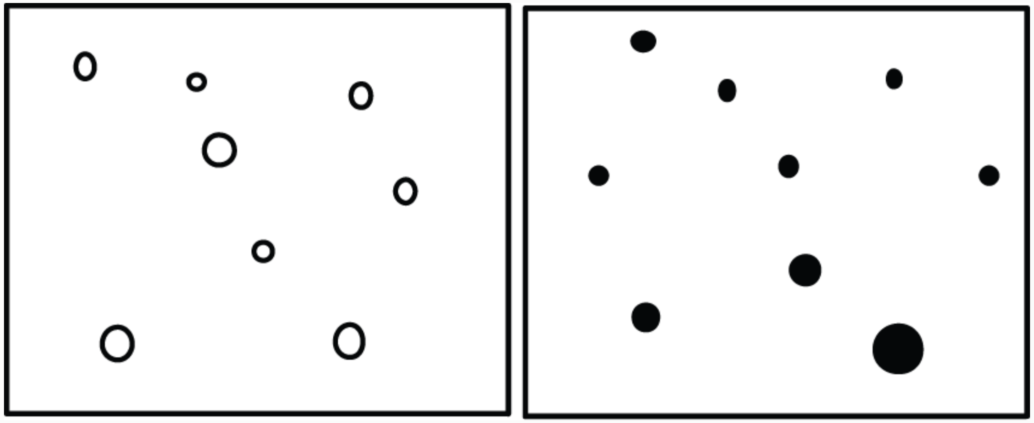
Semantic Substrates



Semantic Substrates

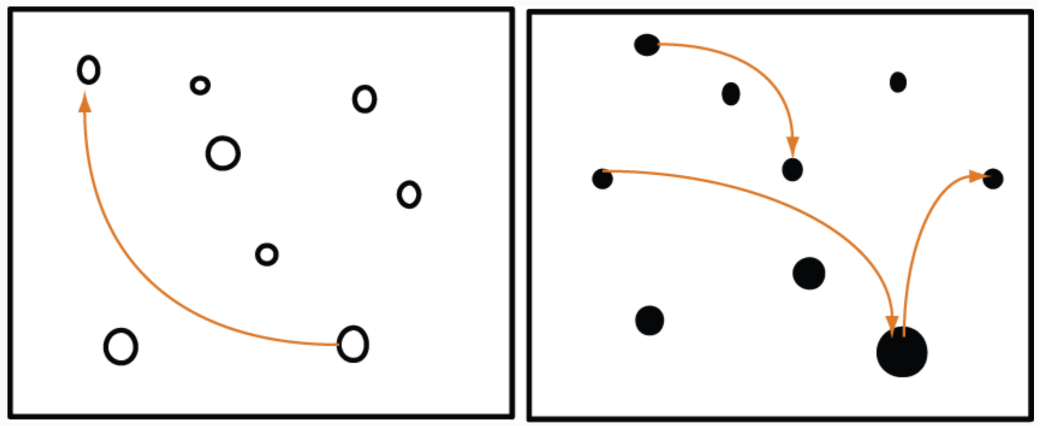


Semantic Substrates



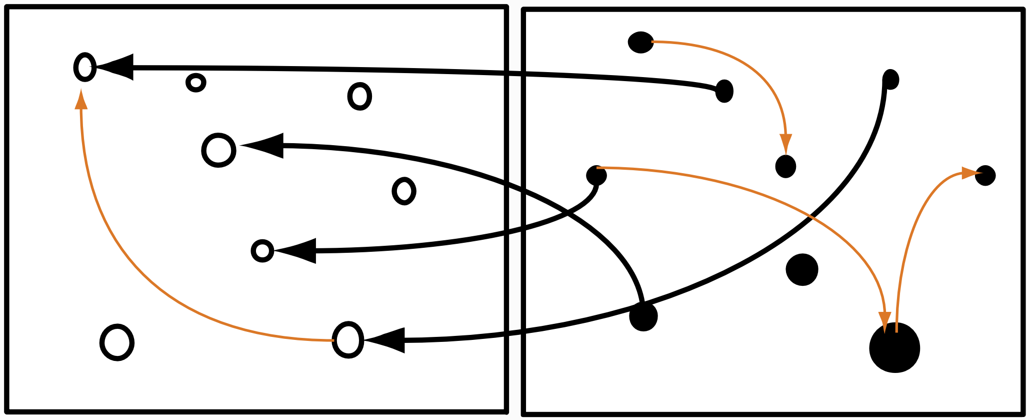
D_{A_1} D_{A_2}

Semantic Substrates



$$Vis_A \rightarrow R_A(D_{A_1}) \qquad Vis_A \rightarrow R_A(D_{A_2})$$

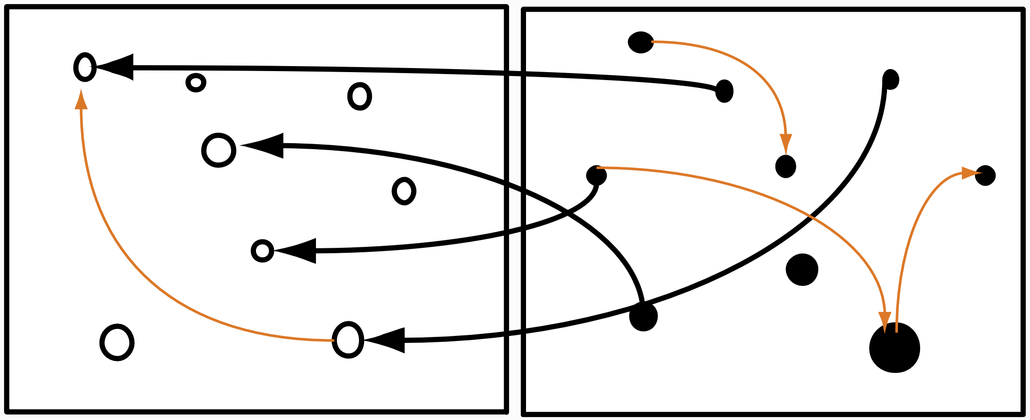
Semantic Substrates



$$Vis_A \rightarrow R_A(D_{A_1}) \qquad Vis_A \rightarrow R_A(D_{A_2})$$

$$Vis_A \rightarrow R_A(D_A)$$

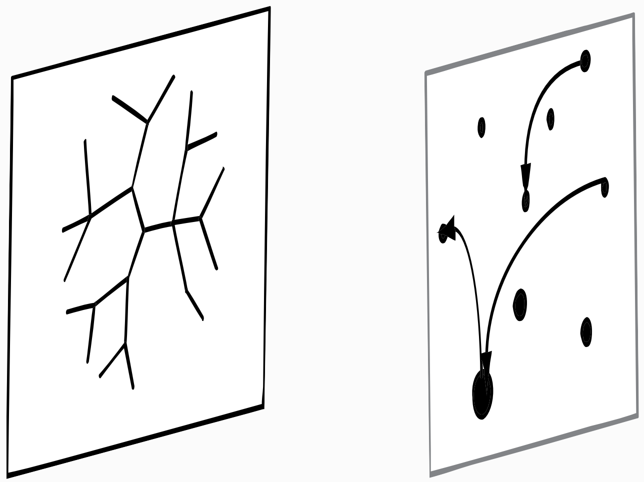
Semantic Substrates



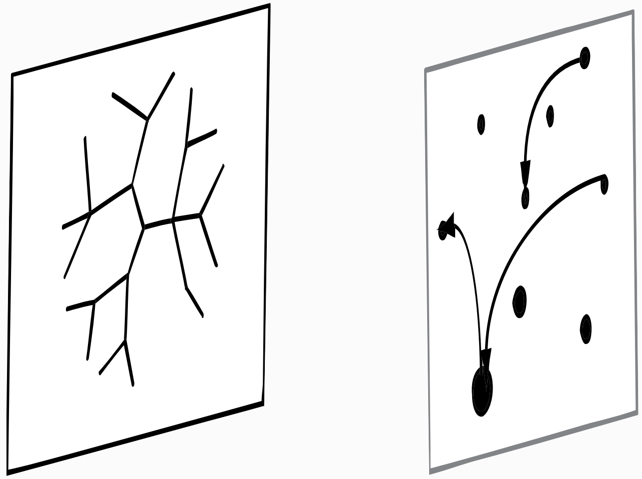
- Single visualization, single relation
- Semantically meaningful data subsets
- Spatial rights for all relations

(Shneiderman and Aris, 2006)

VisLink

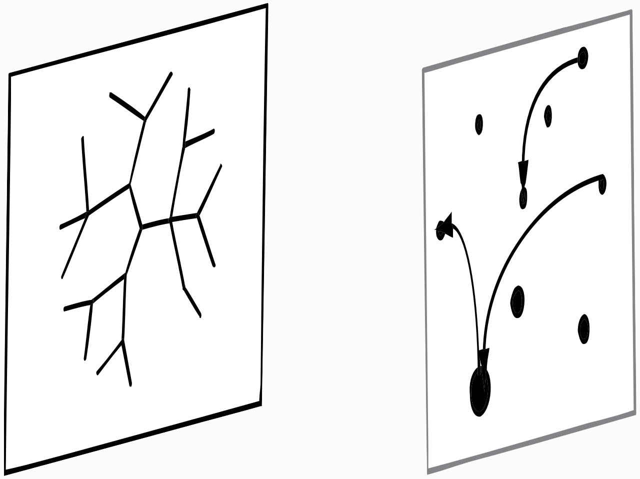


VisLink



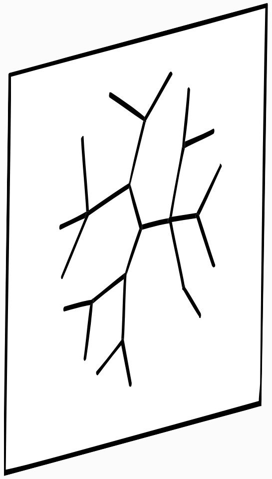
$$Vis_A \rightarrow R_A(D_A)$$

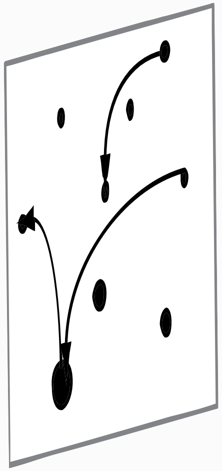
VisLink



$$Vis_A \rightarrow R_A(D_A) \qquad Vis_B \rightarrow R_B(D_A)$$

VisLink



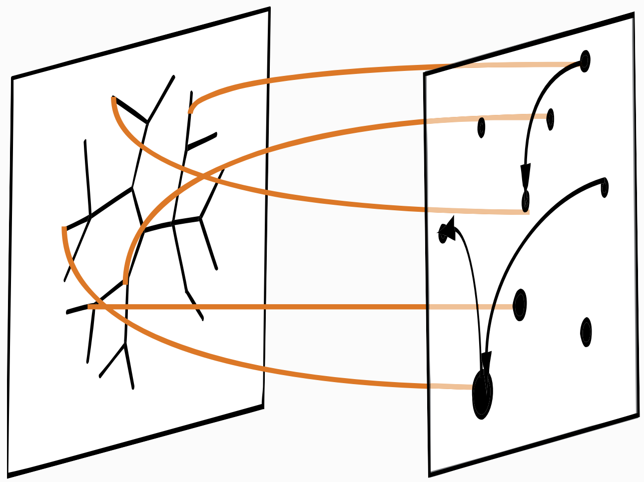


$$Vis_A \rightarrow R_A(D_A)$$

$$Vis_B \rightarrow R_B(D_A)$$

$$Vis_B \rightarrow R_A(D_B)$$

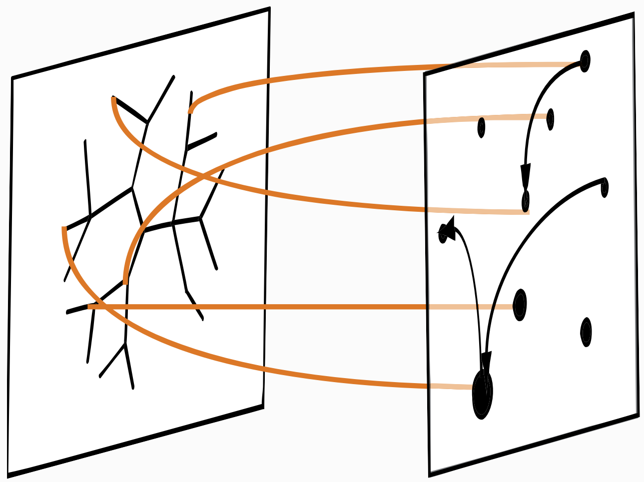
VisLink



$$Vis_A \to R_A(D_A) \qquad Vis_B \to R_B(D_A)$$

$$Vis_{A+B} \to T(R_A(D_A), R_B(D_A))$$

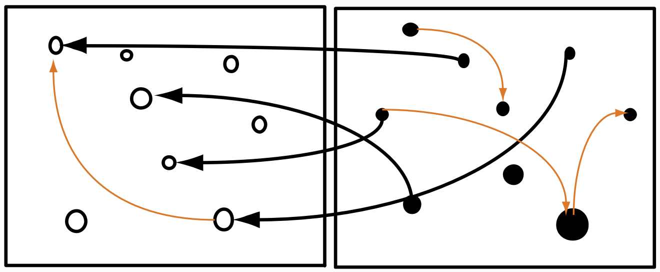
VisLink



$$Vis_{A+B} \rightarrow T(R_A(D_A), R_B(D_A))$$

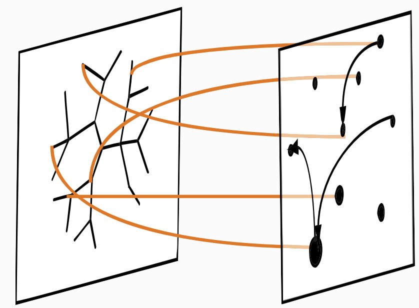
- Visualize second order relations between visualizations
- Across any datasets, relations, visualizations for which a relation can be defined
- All component visualizations retain spatial rights

VisLink & Semantic Substrates



$Vis_A \rightarrow R_A(D_{A_1})$ $\quad$ $Vis_A \rightarrow R_A(D_{A_2})$

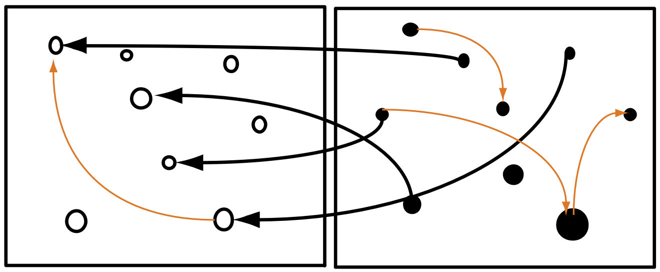
$$\underbrace{Vis_A \rightarrow R_A(D_{A_1}) \quad Vis_A \rightarrow R_A(D_{A_2})}_{Vis_A \rightarrow R_A(D_A)}$$



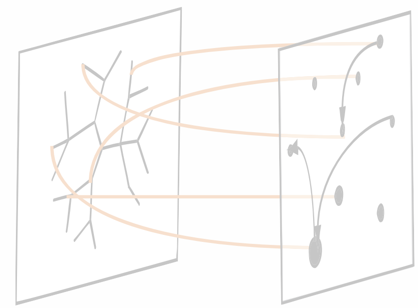
$Vis_A \rightarrow R_A(D_A)$ $\quad$ $Vis_B \rightarrow R_B(D_A)$

$$\underbrace{Vis_A \rightarrow R_A(D_A) \quad Vis_B \rightarrow R_B(D_A)}_{Vis_{A+B} \rightarrow T(R_A(D_A), R_B(D_A))}$$

VisLink & Semantic Substrates



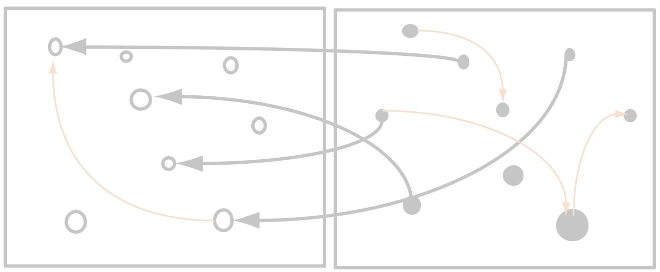
$$\underbrace{Vis_A \to R_A(D_{A_1}) \qquad Vis_A \to R_A(D_{A_2})}_{Vis_A \to R_A(D_A)}$$



$$\underbrace{Vis_A \to R_A(D_A) \qquad Vis_B \to R_B(D_A)}_{Vis_{A+B} \to T(R_A(D_A), R_B(D_A))}$$

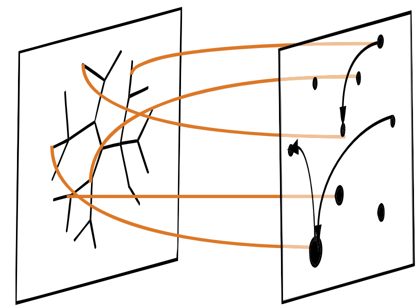
- Single visualization technique
- Semantic subsets of data provide added meaning

VisLink & Semantic Substrates



$Vis_A \rightarrow R_A(D_{A_1})$ $Vis_A \rightarrow R_A(D_{A_2})$

$Vis_A \rightarrow R_A(D_A)$

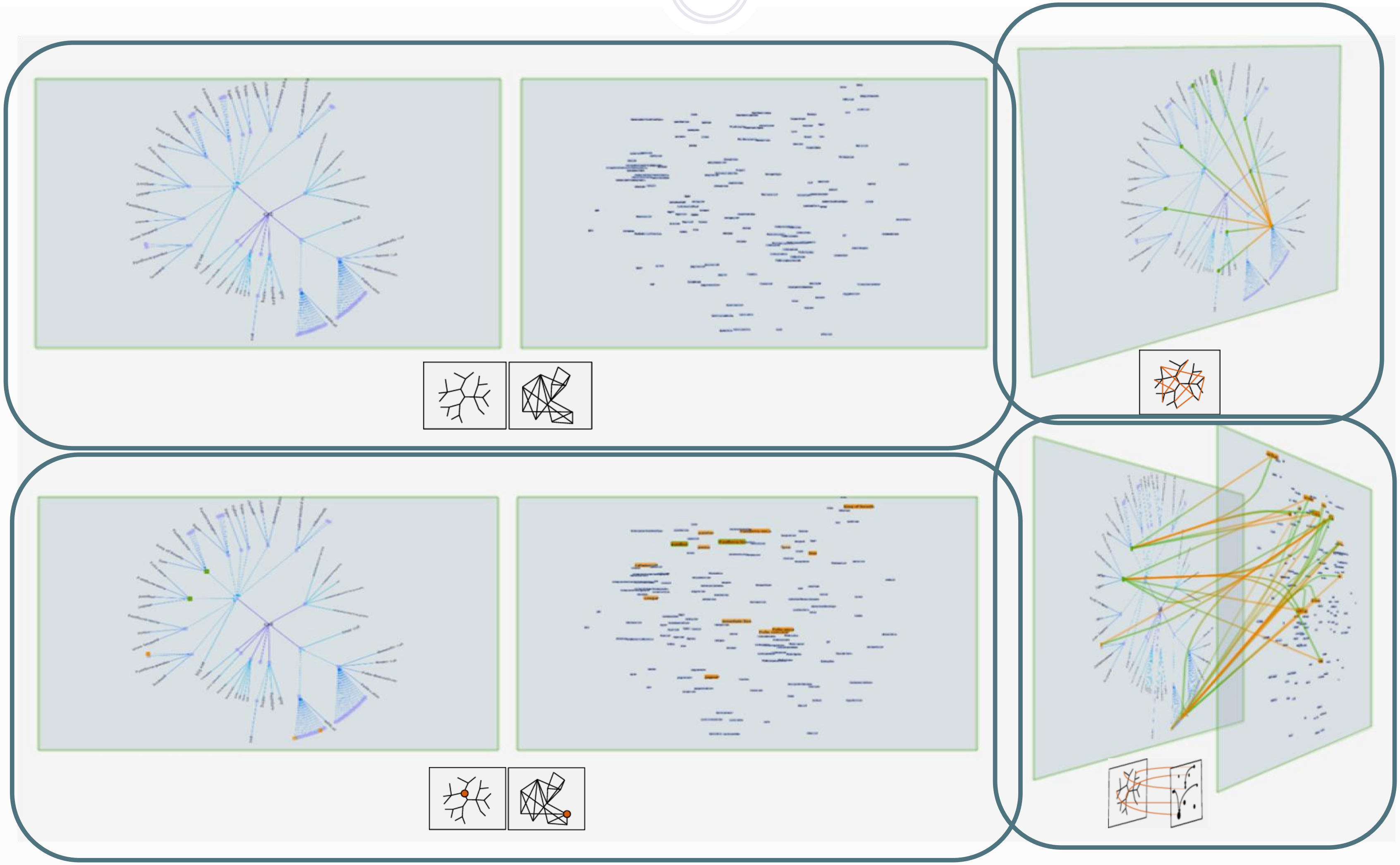


$Vis_A \rightarrow R_A(D_A)$ $Vis_B \rightarrow R_B(D_A)$

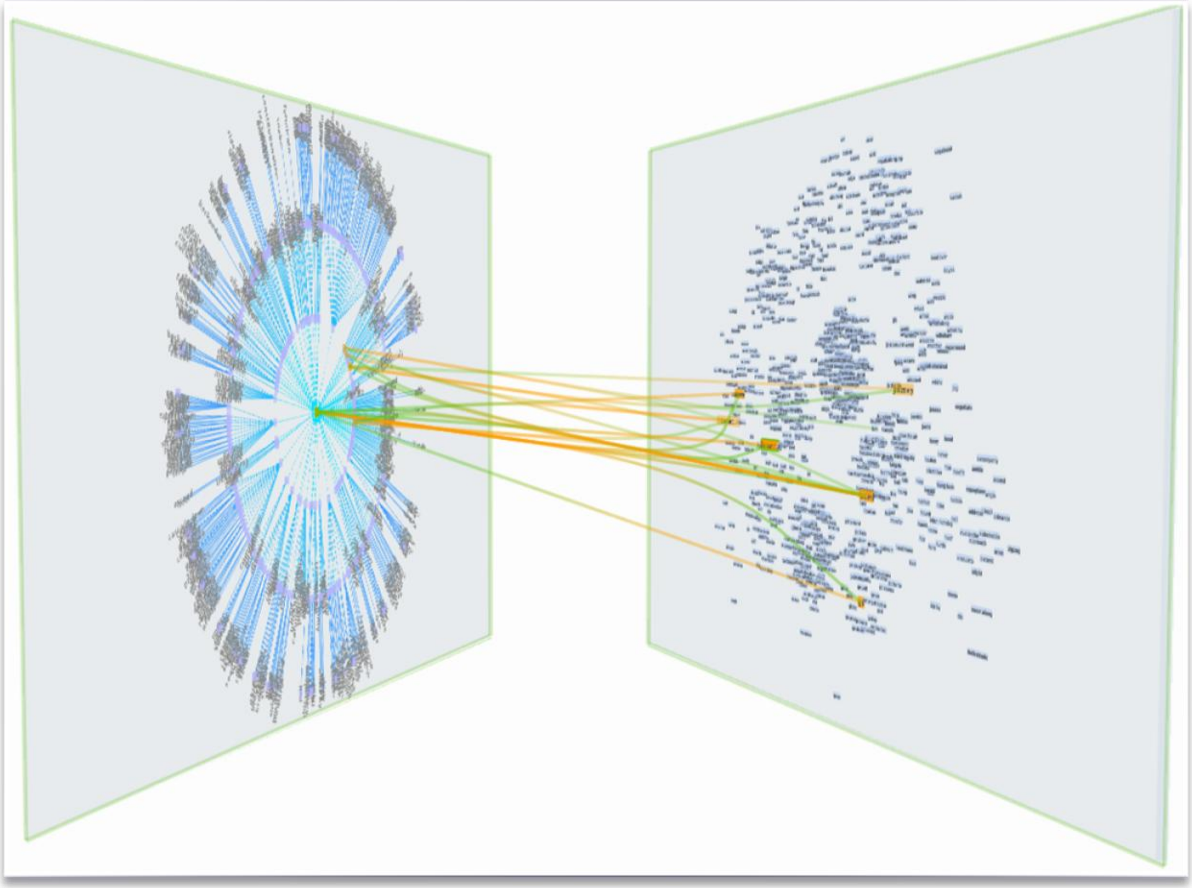
$Vis_{A+B} \rightarrow T(R_A(D_A), R_B(D_A))$

- Any number of different relations and visualizations
- Second order relations revealed in inter-plane edges

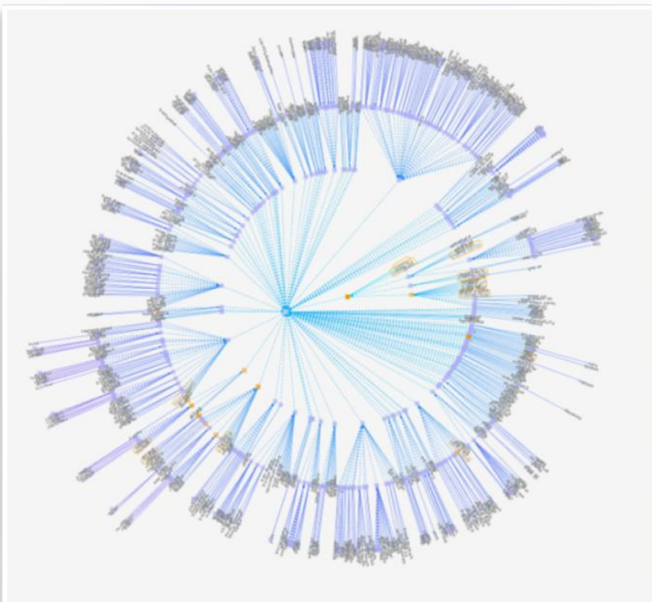
Equivalency & Extension

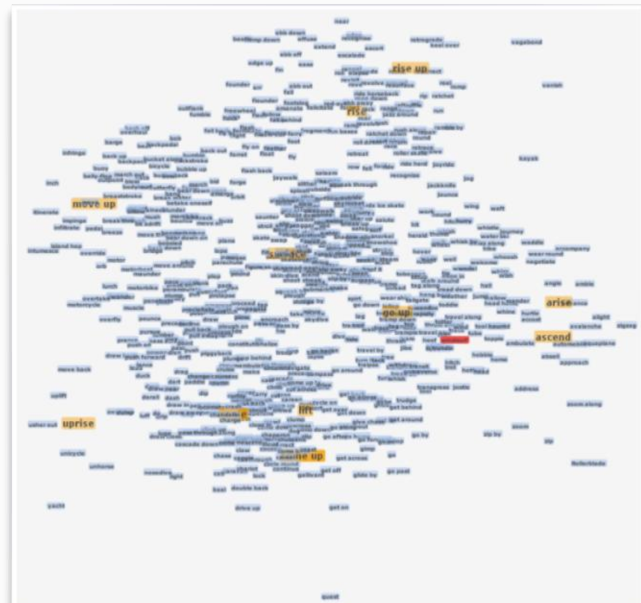


VisLink



VisLink Case Study: Lexical Data



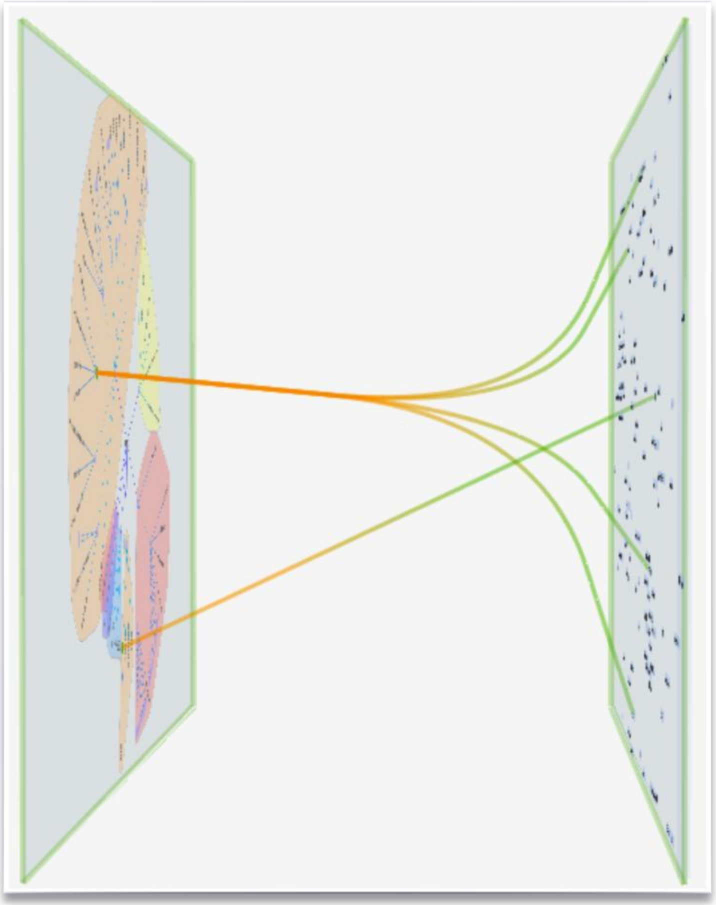


?

WordNet IS-A hierarchy (R_A) using radial tree (Vis_A)

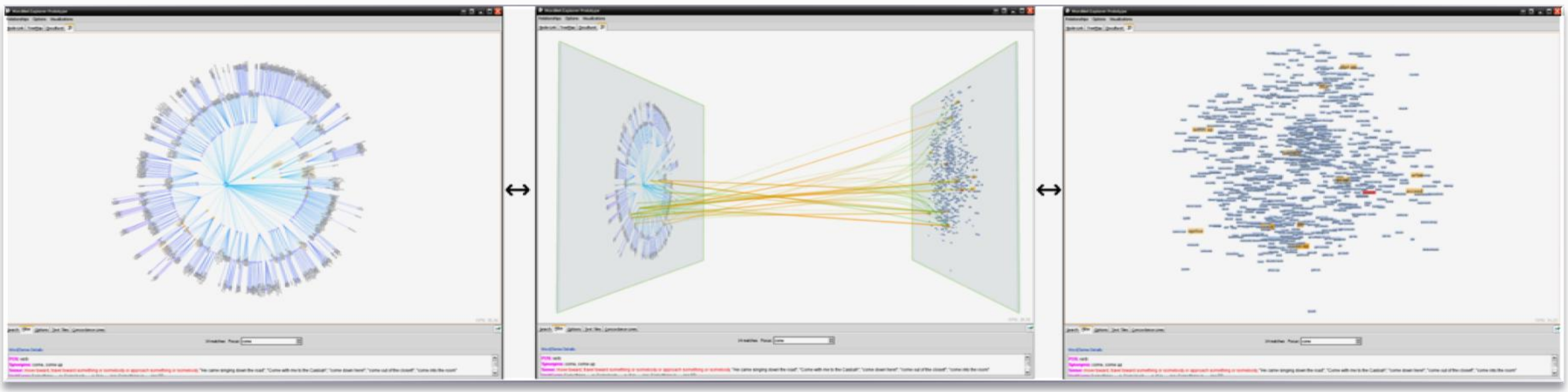
Similarity clustering (R_B) using force-directed layout (Vis_B)

Edge Detail



- Bundled:
  one-to-many edges
- Smooth:
  Chaiken corner cutting
- Transparent:
  bundles more opaque
- Directed:
  orange-to-green

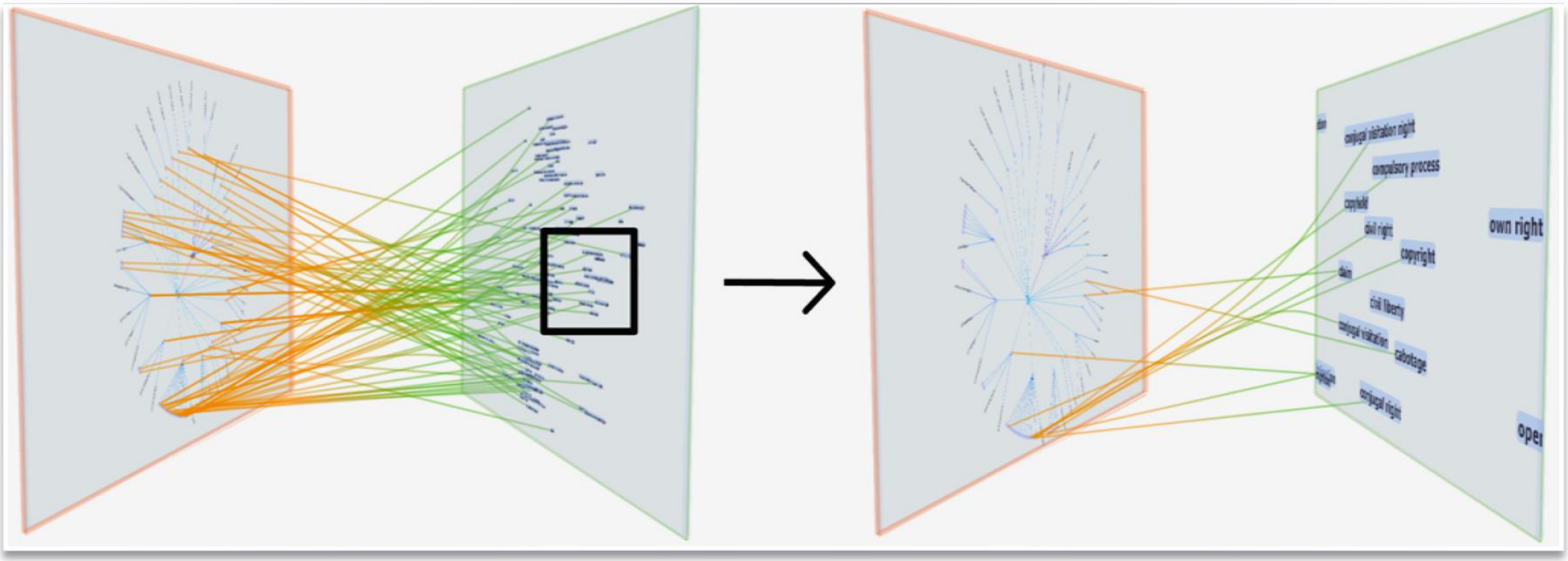
Interaction With Component Visualizations



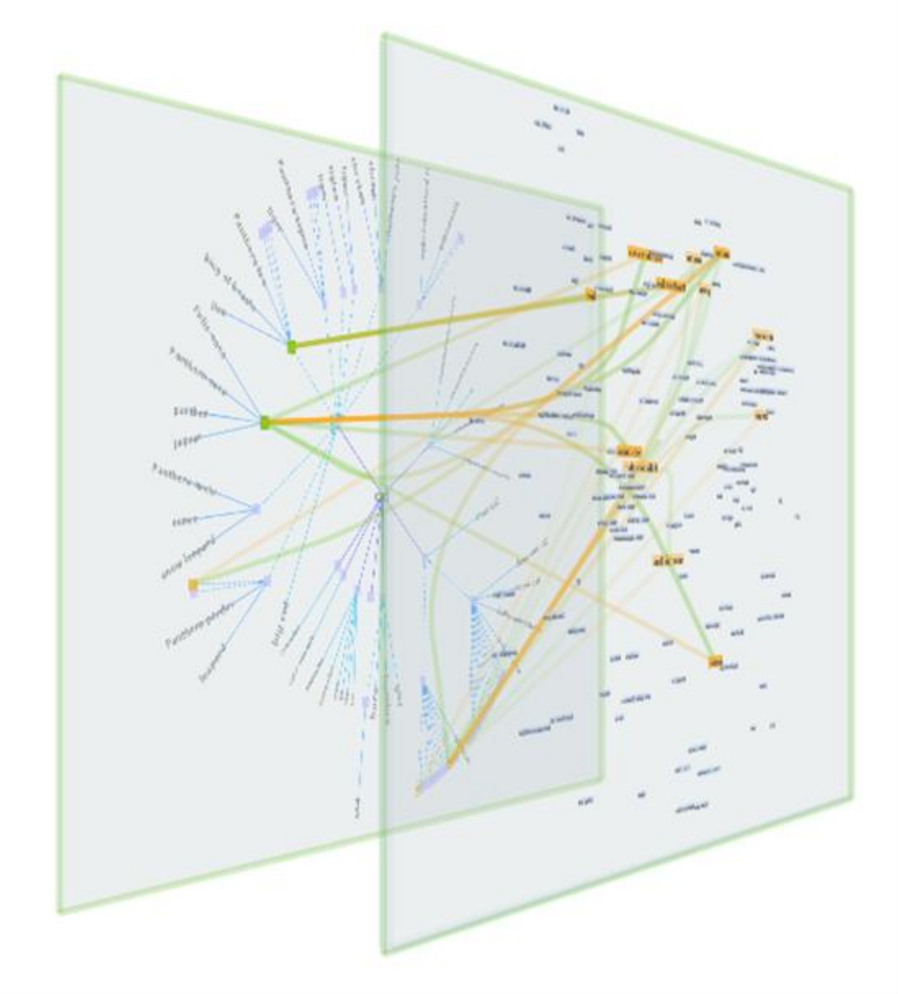
- Always equivalent to 2D:
  - Planes are virtual displays
  - Mouse events transformed and passed to underlying visualization
  - Equivalent to 2D viewing mode

Interplane Edges

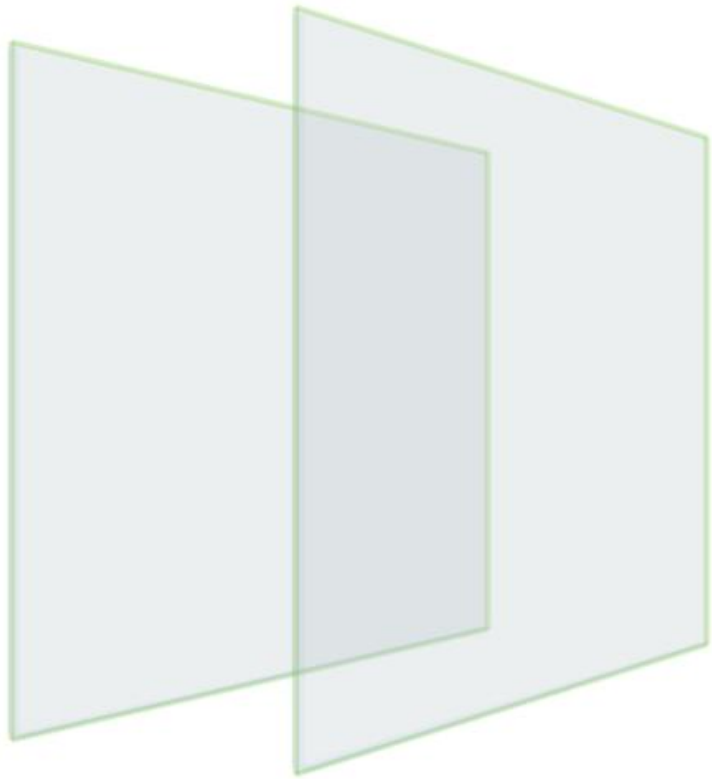
Zoom


conjugal visitation right
compulsory process
own right
civil right
copyright
civil liberty
conjugal visitation
cabotage
conjugal right

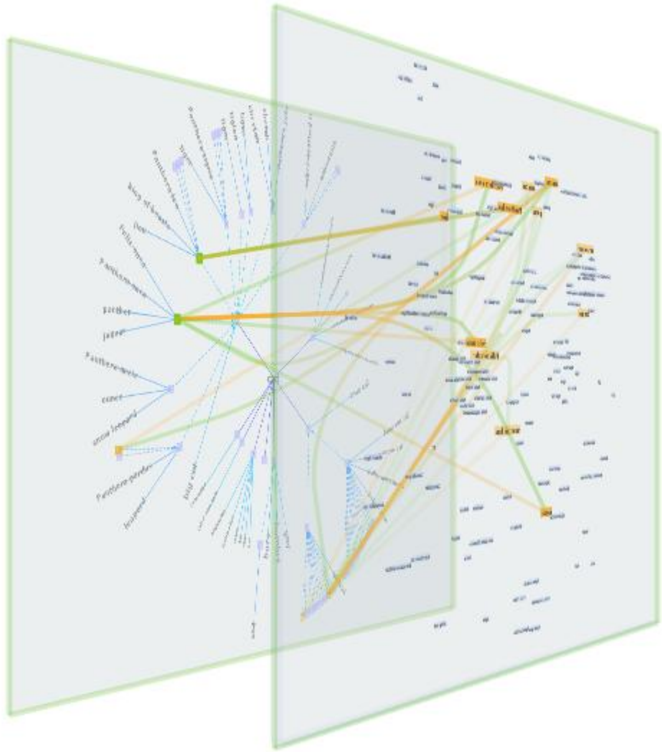
Filter



Constrained Widget Interaction



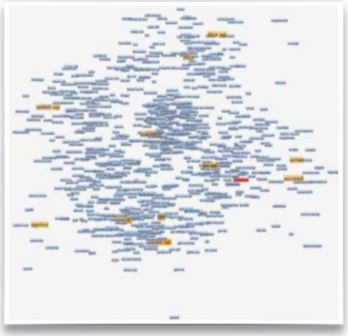
3D Navigation



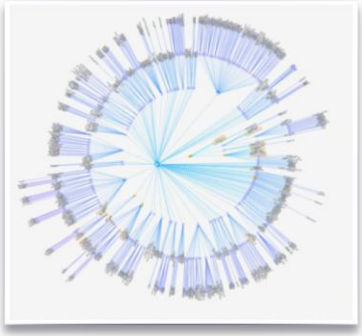
Spreading Activation

- Nodes have a level of activation, indicated by transparency of orange node background
- Full activation through:
  - Selecting a node on a plane
  - Node matches search query
- Activation propagates through interplane edges, reflecting between planes with exponential drop-off
- Enables inter-visualization queries
- Edge transparency relative to source node activation

Inter-Plane Query Example



1: alphabetic clusters



2: synonym sets

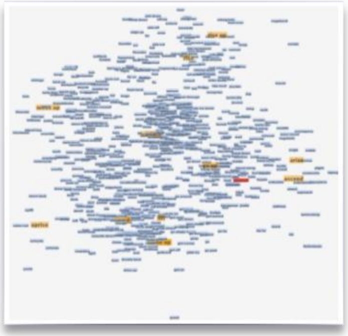
No synonym information

No alphabetic organization

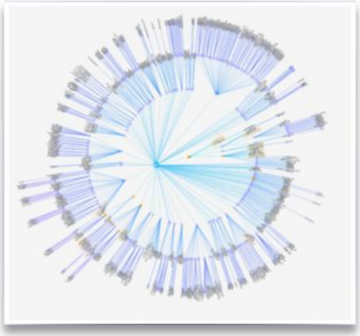
Q: Synonyms in the alphabetic view?

Inter-Plane Query Example

1. Select a word on plane 1
2. Edges propagate to synonym sets on plane 2
3. Reflected edges propagate back, revealing synonyms in alphabetic clusters



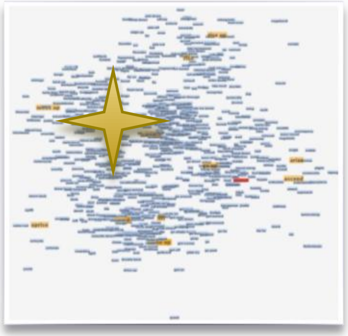
1: similarity clusters



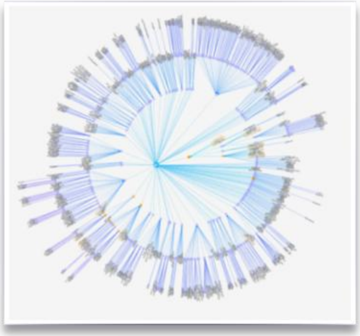
2: synonym sets

Inter-Plane Query Example

1. Select a word on plane 1
2. Edges propagate to synonym sets on plane 2
3. Reflected edges propagate back, revealing synonyms in alphabetic clusters



1: similarity clusters



2: synonym sets

Inter-Plane Query Example

1. Select a word on plane 1
2. Edges propagate to synonym sets on plane 2
3. Reflected edges propagate back, revealing synonyms in alphabetic clusters

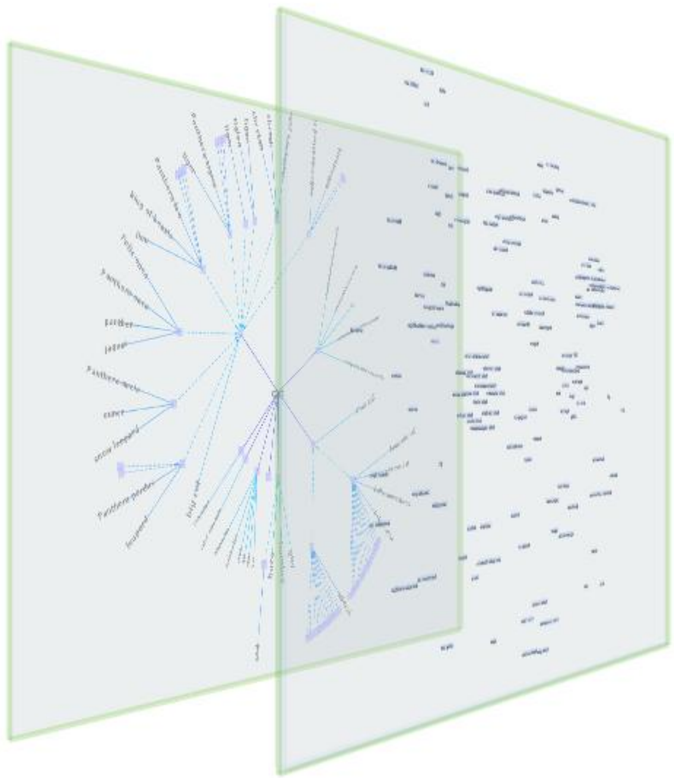
1: similarity clusters 2: synonym sets

Inter-Plane Query Example

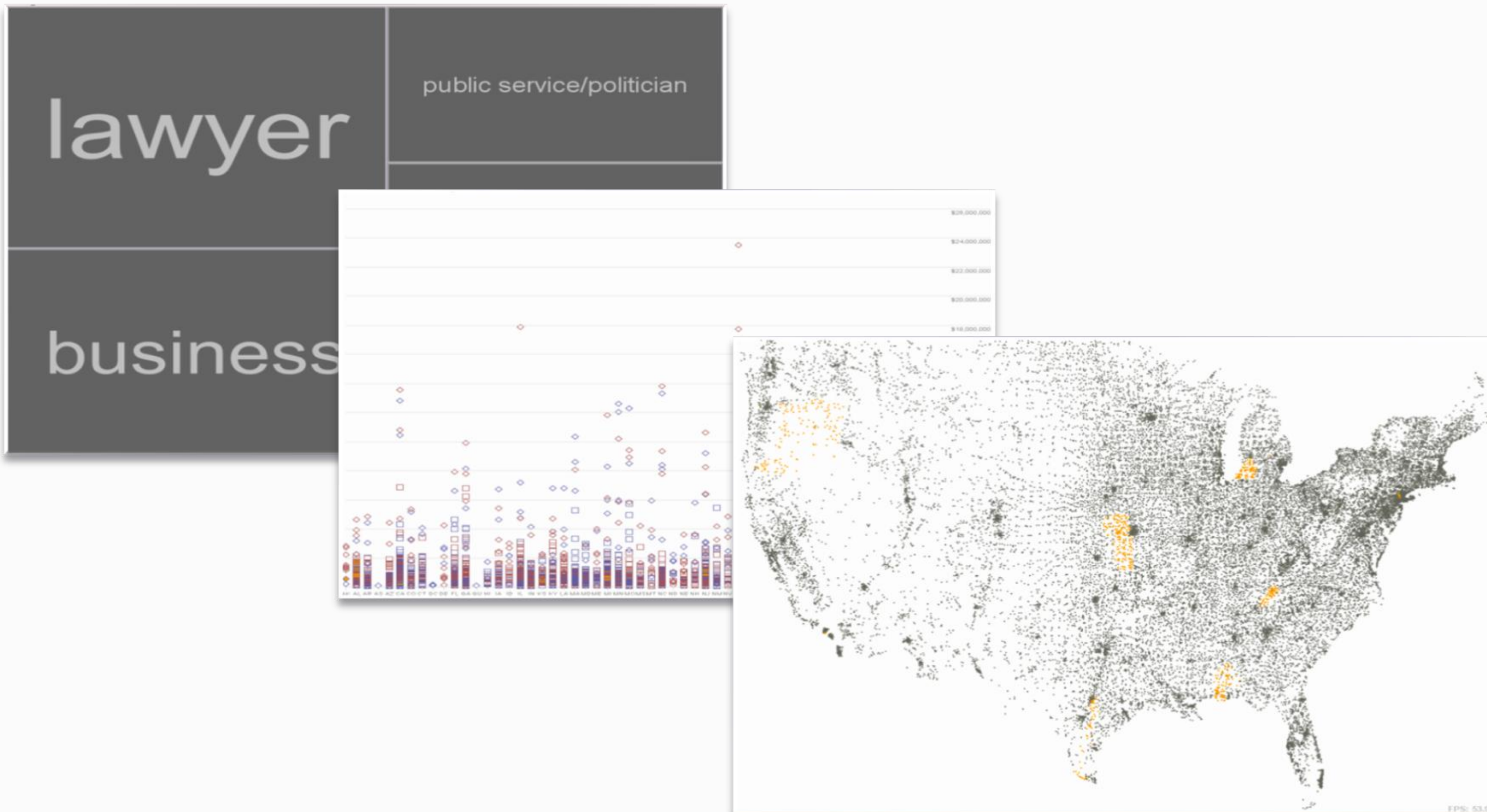
1. Select a word on plane 1
2. Edges propagate to synonym sets on plane 2
3. Reflected edges propagate back, revealing synonyms in alphabetic clusters

1: similarity clusters 2: synonym sets

Edge Reflection and Inter-Plane Queries



Linking Existing Visualizations



Case Study

Linking Existing Visualizations

Implementation

- Prefuse visualization toolkit (Heer et al., 2005)
  - Existing visualizations can be incorporated without changes
  - Interplane edges defined by (plane, node) index pairs
- Java OpenGL

Perceptual Considerations

- Not all layouts equal
- Colour interactions with edges and visualizations
- 3D perspective bias

Future Work

- Application to additional analytic scenarios
- Investigation of 3D edge bundling, edge lenses
- Animation of spreading activation
- Evaluation against existing multiple view techniques
- Rich query language to filter visualization planes

Summary

- Formalism to describe multi-relation visualizations
- New way to reveal relationships amongst visualizations
- Reuse of the powerful spatial visual dimension
- Full 2D interactivity for constituent visualizations
- Techniques to simplify 3D navigation
- Visualization bridging through inter-representational queries and spreading activation

VisLink: Revealing Relationships Amongst Visualizations

Acknowledgements:

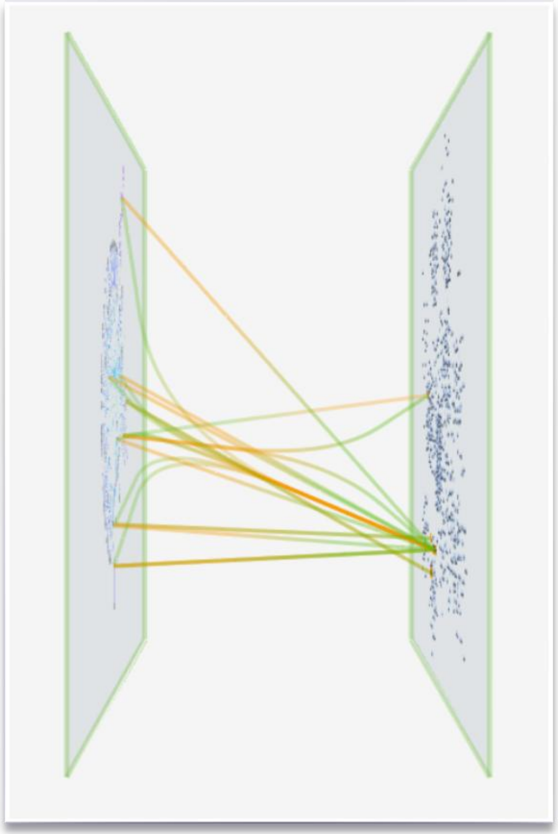
Gerald Penn, Petra Isenberg, Mark Hancock, Tobias Isenberg, Uta Hinrichs, and Matthew Tobiasz, and helpful reviewers for their advice.

NECTAR

NSERC CRSNG

iCORE

SMART Technologies Inc.



Christopher Collins (ccollins@cs.utoronto.ca) & Sheelagh Carpendale (sheelagh@ucalgary.ca)