



# Visual Analysis of Set Relations in a Graph

**Panpan Xu<sup>1</sup>**, Fan Du<sup>2</sup>, Nan Cao<sup>3</sup>, Conglei Shi<sup>1</sup>, Hong Zhou<sup>4</sup>, Huamin Qu<sup>1</sup>

<sup>1</sup> Hong Kong University of Science and Technology, 2 Zhejiang University,

3 IBM T. J. Watson Research Center, 4 Shenzhen University

2013.06.19

Motivation: data model and research questions

Approaches

Previous works

Technical details

Case studies

Limitation and future works

Outline

**Motivation: data model and research questions**

Approaches

Previous works

Technical details

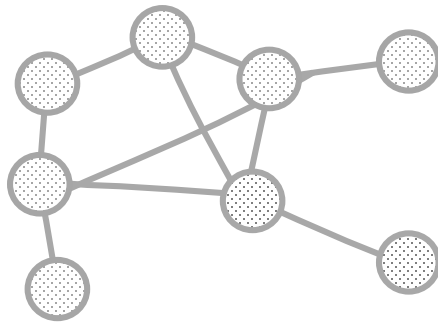
Case studies

Limitation and future works

Outline



## Collaboration network



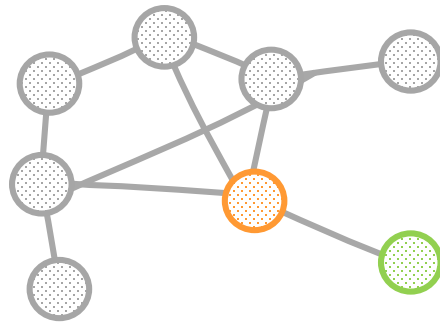
## Research topics

- tree
- graph
- hierarchical data
- • pipeline
- architecture

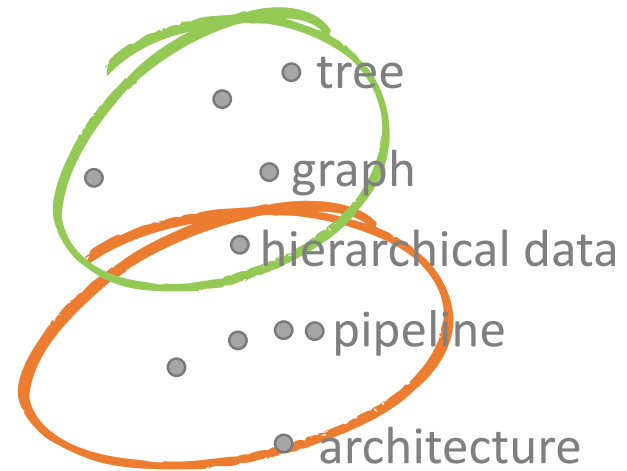
Data



# Collaboration network



# Research topics



Data



# Homophily effect

Research questions



# Homophily effect

Do “birds of a feather flock together”?

Research questions



# Homophily effect

Do “birds of a feather flock together”?

How proximity of nodes correlates to set relation?

Research questions





## Homophily effect

Do “birds of a feather flock together”?

How proximity of nodes correlates to set relation?

## Set relation over item clusters

Distribution and implicit overlap of the sets

Research questions



*complementary  
perspectives*

## Homophily effect

Do “birds of a feather flock together”?

How proximity of nodes correlates to set relation?

## Set relation over item clusters

Distribution and implicit overlap of the sets

Research questions

Motivation: data model and research questions

## **Approaches**

Previous works

Technical details

Case studies

Limitation and future works

## Outline



Homophily effect

Set relation over item clusters

Approaches



# Homophily effect

**Glyph design** at graph nodes correlates set relation and node distance

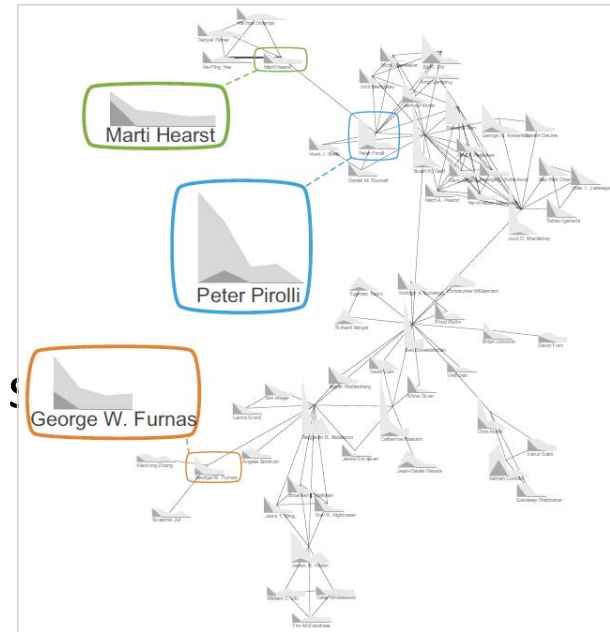
Set relation over item clusters

Approaches

# Homophily effect

**Glyph design** at graph nodes correlates with node distance

Set relation over item clusters



# Approaches



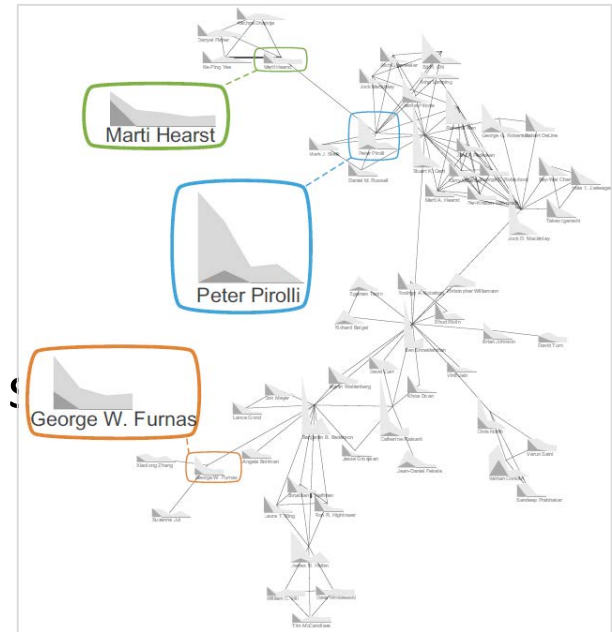
# Homophily effect

**Glyph design** at graph nodes correlates with node distance

# Set relation over item clusters

**Contour map + visual link** design

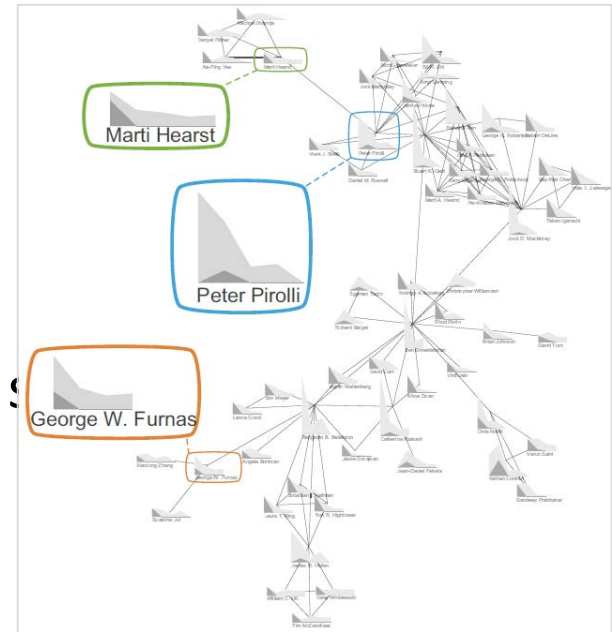
**Layout algorithm** trades precise location of the items for visual simplicity (inspired by metro map drawing, storyline visualization)



# Approaches

# Homophily effect

**Glyph design** at graph nodes correlates with node distance



# Set relation over item clusters

**Contour map + visual link** design

**Layout algorithm** trades precise local placement for overall simplicity (inspired by metro map design)



Approaches



Motivation: data model and research questions

Approaches

**Previous works**

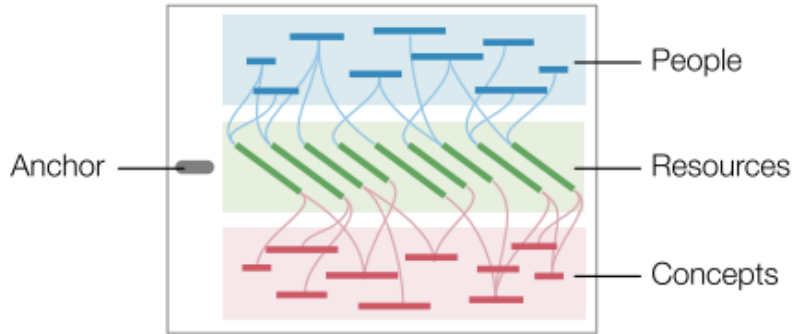
Technical details

Case studies

Limitation and future works

Outline

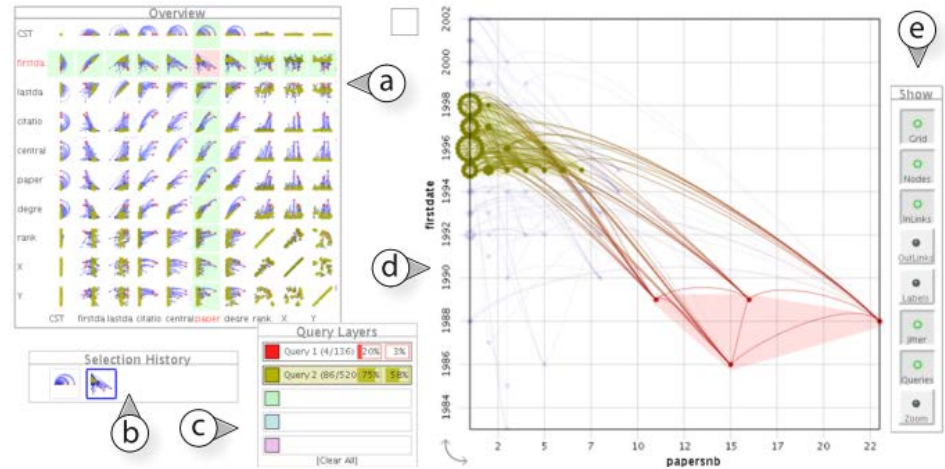




PivotPath [Dörk et al. 12]

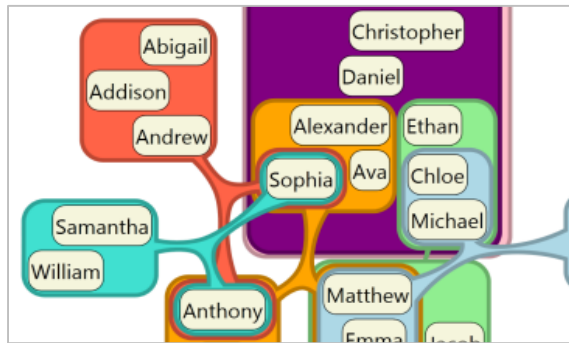


FacetAtlas [Cao et al. 10]



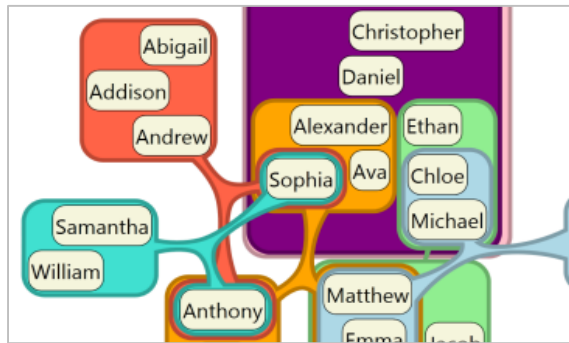
GraphDice [Bezerianos et al. 10]

Previous works - graph visualization

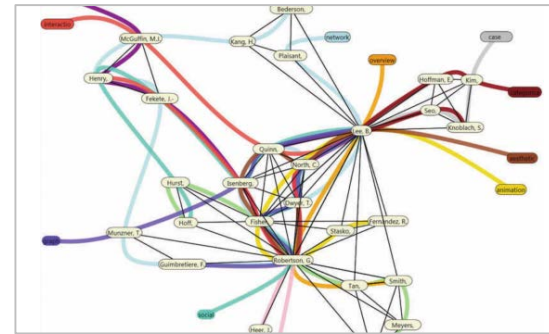


Untangling Euler diagrams  
[Riche and Dwyer, 10]

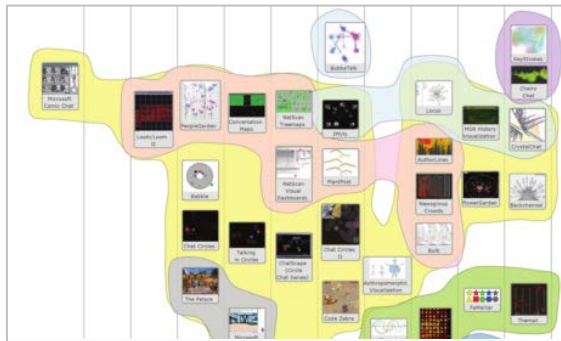
Previous works - set visualization



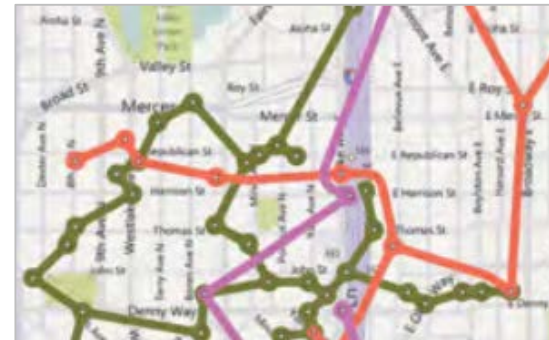
Untangling Euler diagrams  
[Riche and Dwyer, 10]



Line Set [Alper et al. , 11]



Bubble Set [Collins et al., 09]



Kelp Diagram [Dinkla et al., 12]

Previous works – set visualization

Motivation: data model and research questions

Approaches

Previous works

**Technical details**

Case studies

Limitation and future works

Outline



Correlate set overlap and node distance

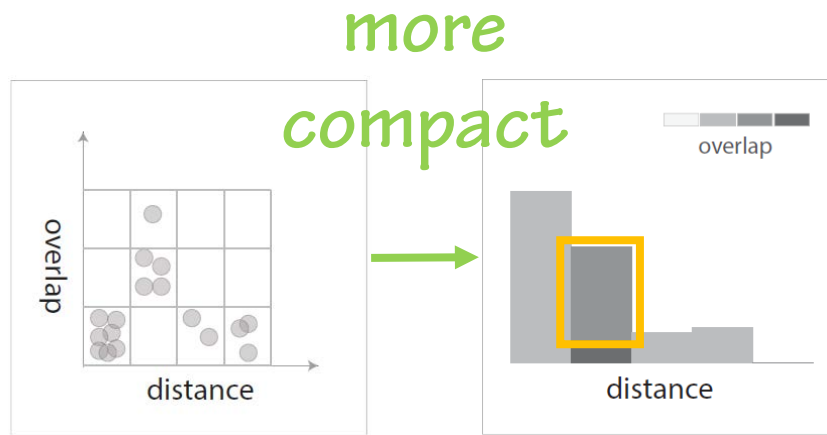


Scatterplot

Glyph design



# Correlate set overlap and node distance



**Shade**  $\propto$  amount of set overlap  
**Height**  $\propto$  the number of nodes at same distances and with similar amount of overlap

Scatterplot

Stacked Barchart

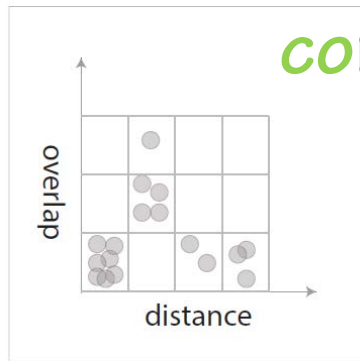
Glyph design



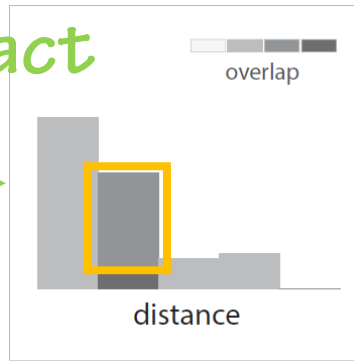


# Correlate set overlap and node distance

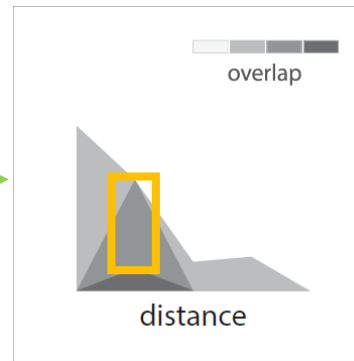
more  
compact



Scatterplot

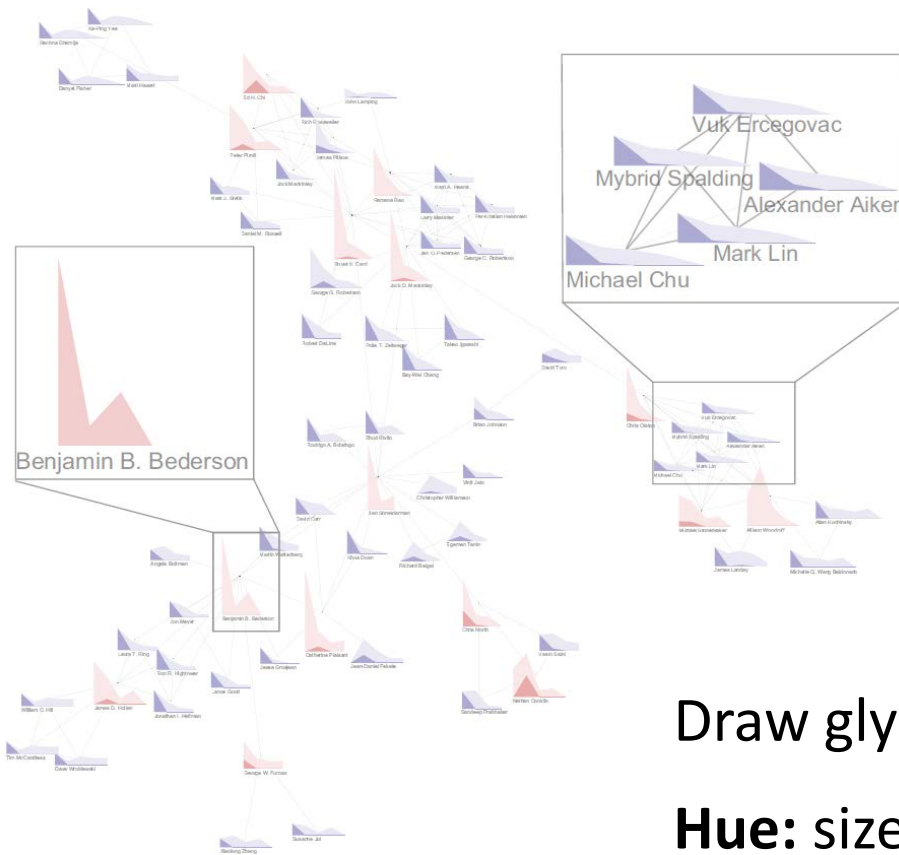


Stacked Barchart



Stacked Graph

Glyph design



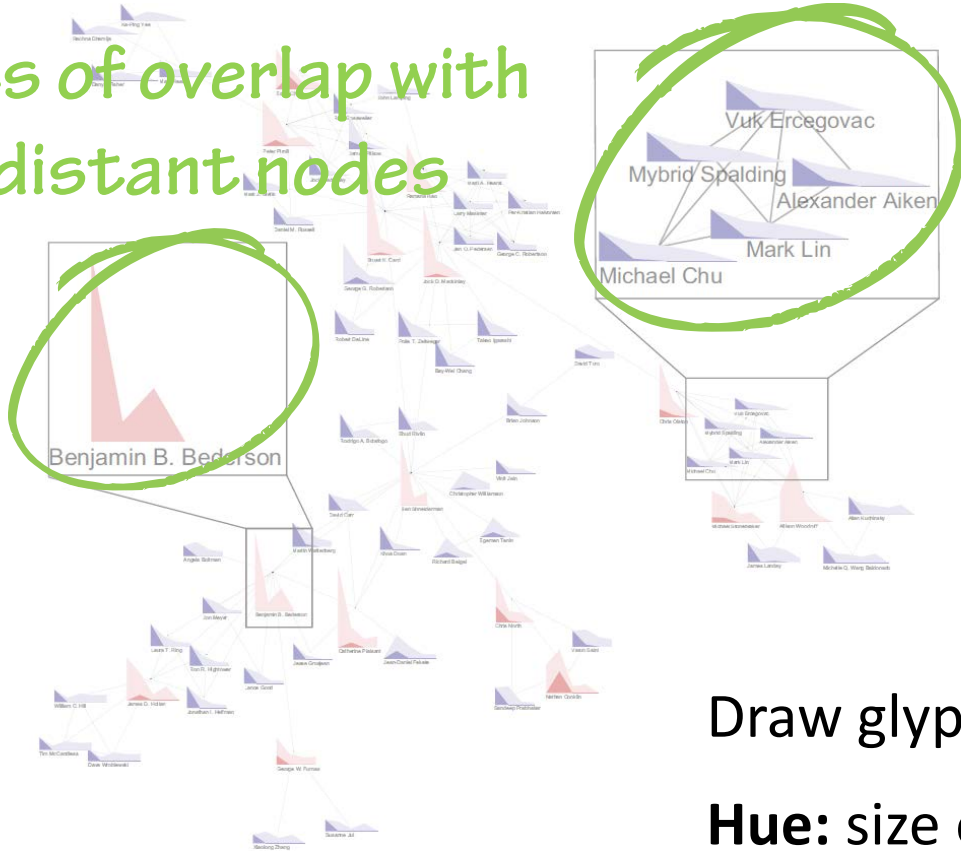
Draw glyphs for each node on a graph  
**Hue:** size of the set compared to its neighbors

# Glyph design



lots of overlap with distant nodes

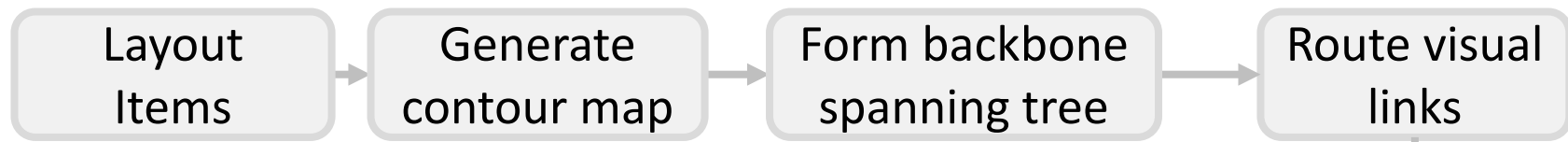
community with locally distributed interests



Draw glyphs for each node on a graph

**Hue:** size of the set compared to its neighbors

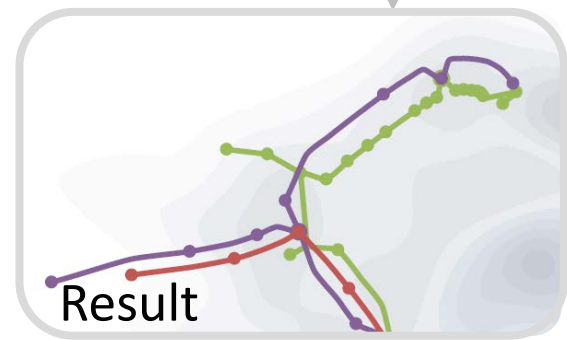
Glyph design

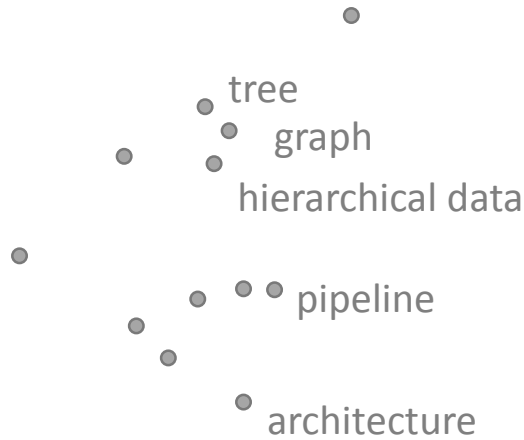


Visually summarize item clusters

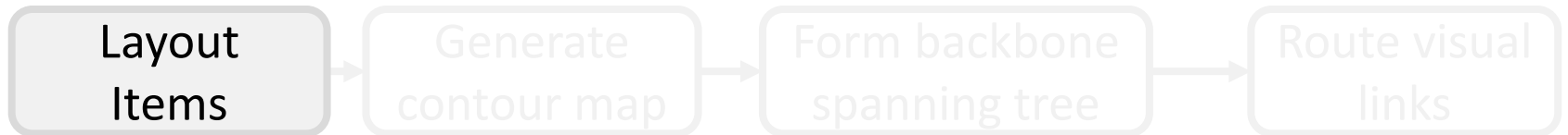
Layout visual links for sets

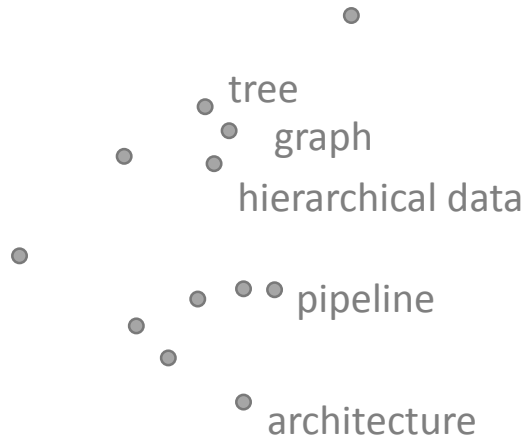
Set visualization over item clusters



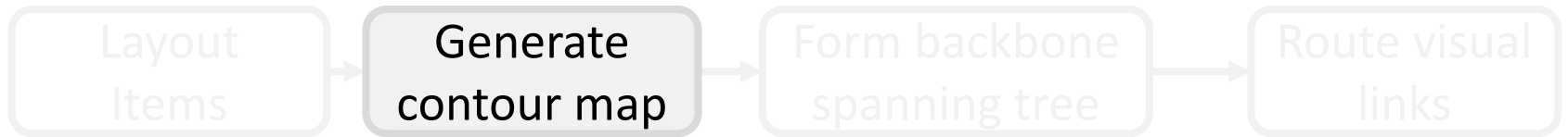


MDS: similar items form visual clusters



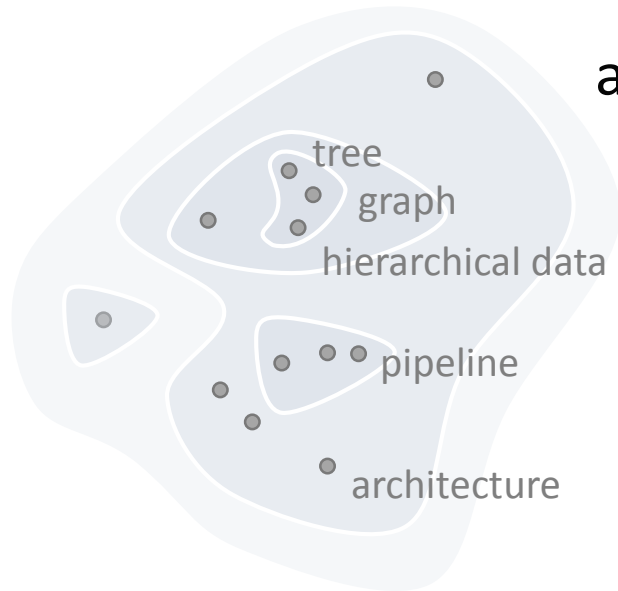


MDS: similar items form visual clusters

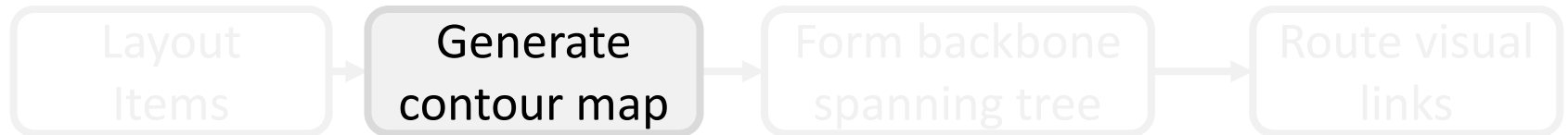




Contour map with KDE:  
abstracted display of item clusters

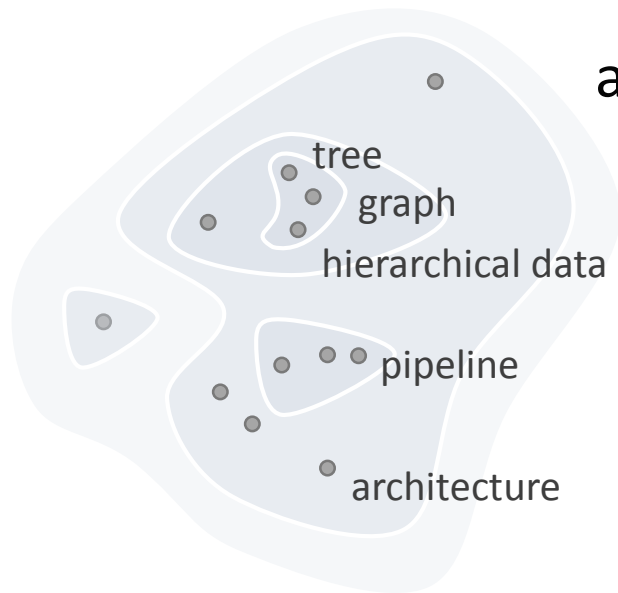


MDS: similar items form visual clusters



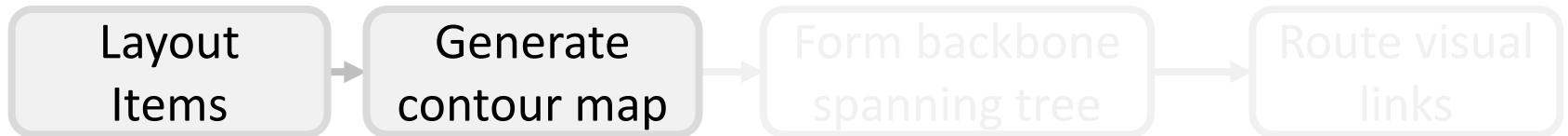


Contour map with KDE:  
abstracted display of item clusters

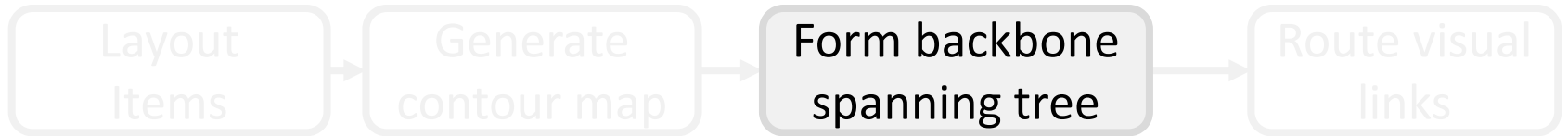
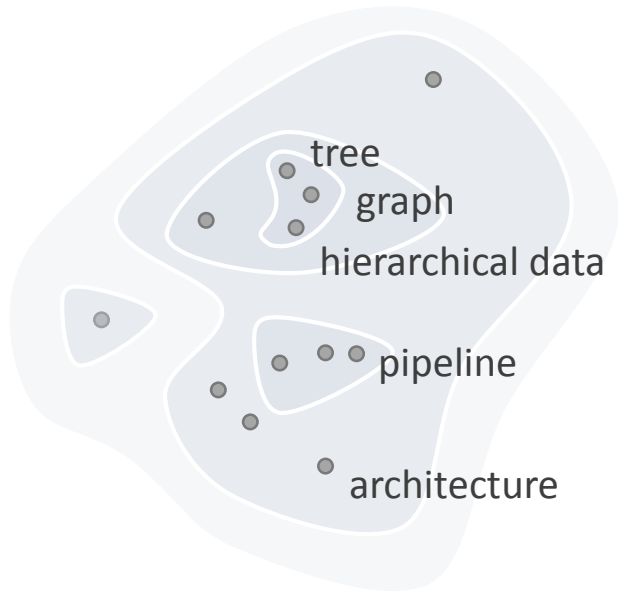


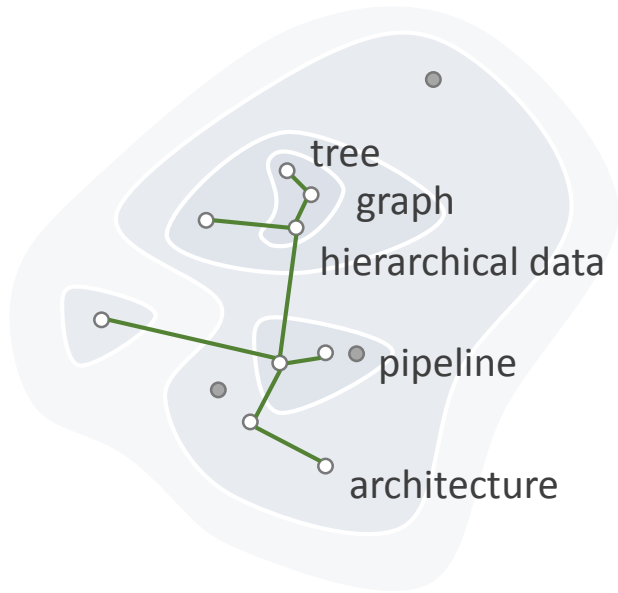
*form context for drawing the sets*

MDS: similar items form visual clusters

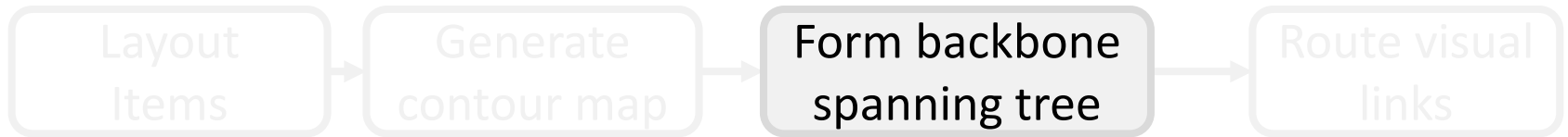


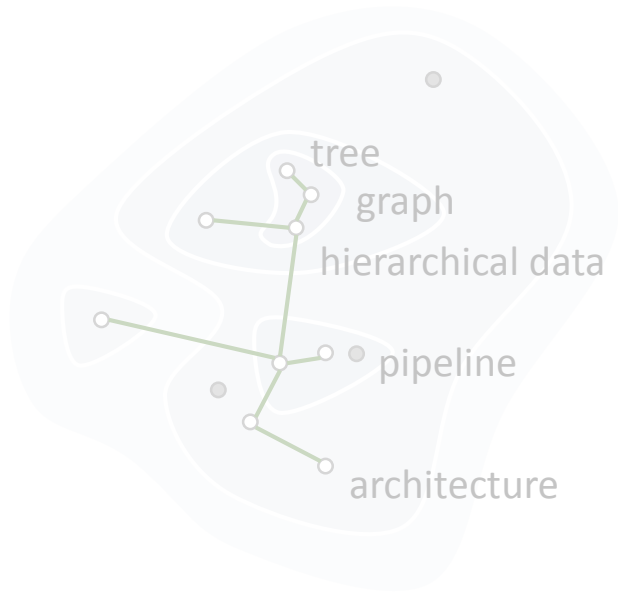




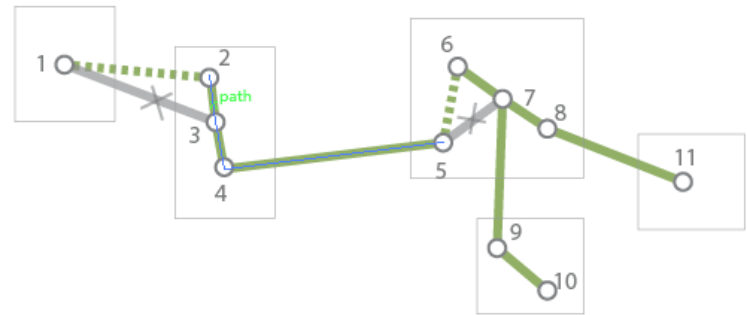


Form MST for items  
in selected sets

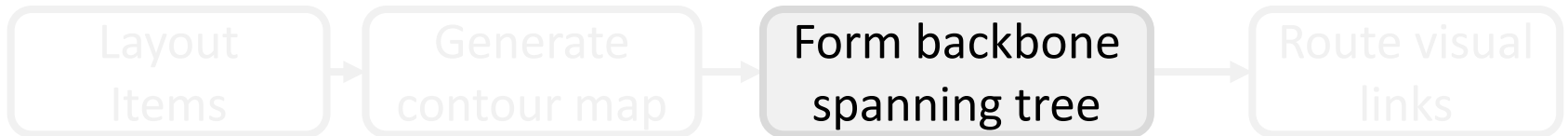


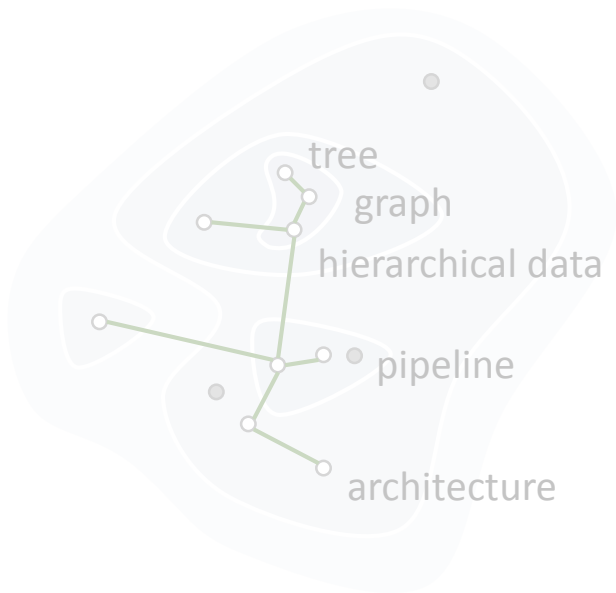


Form MST for items  
in selected sets

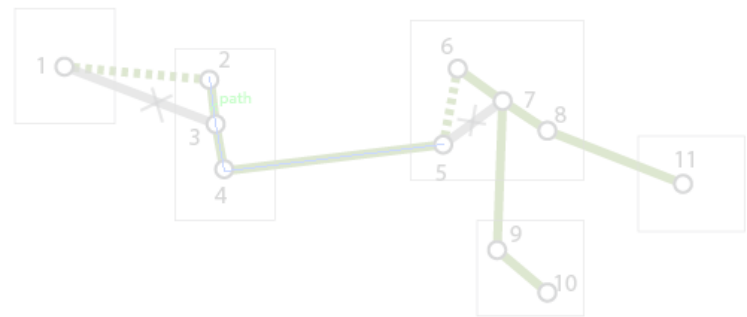


Fold small branches on MST

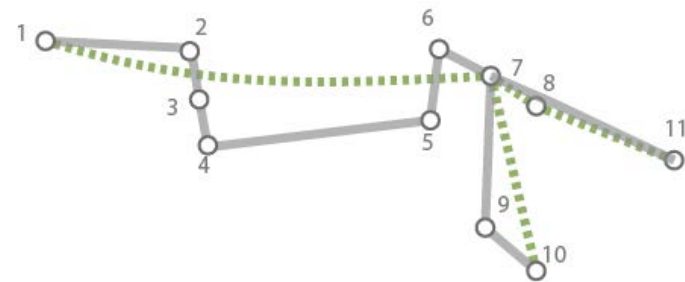




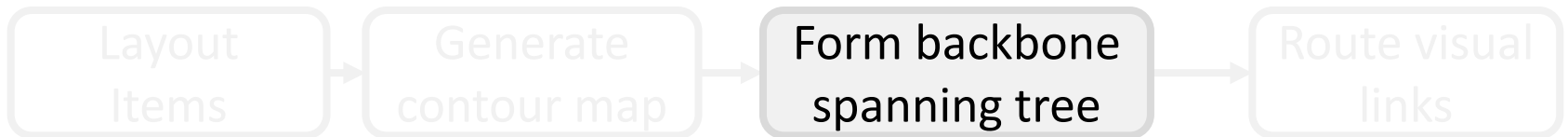
Form MST for items  
in selected sets

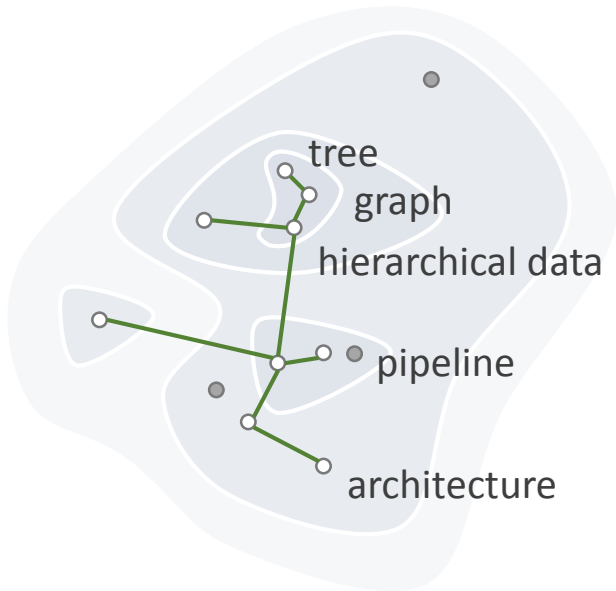


Fold small branches on MST

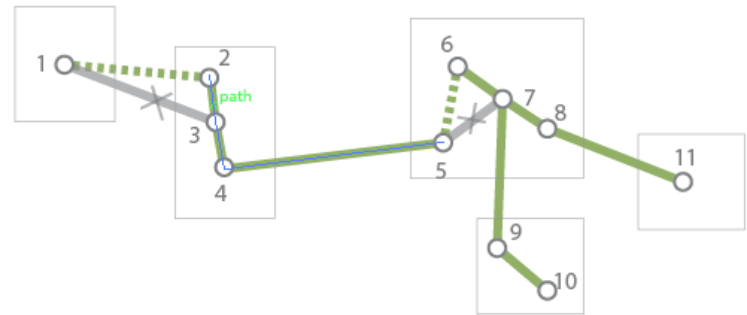


Straighten branches

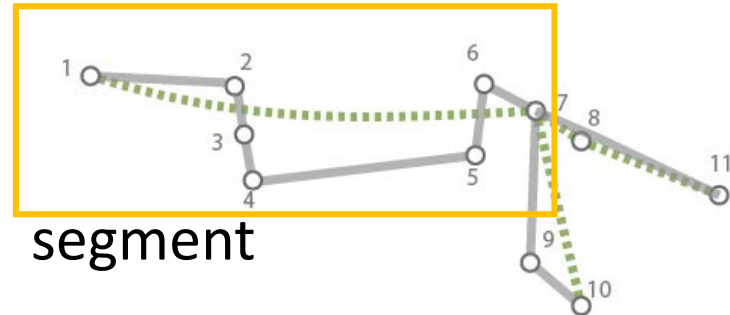




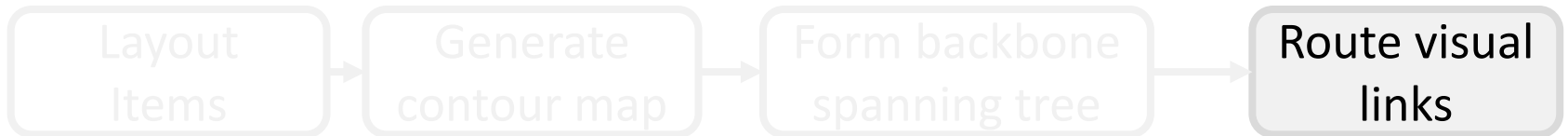
Form MST for items  
in selected sets

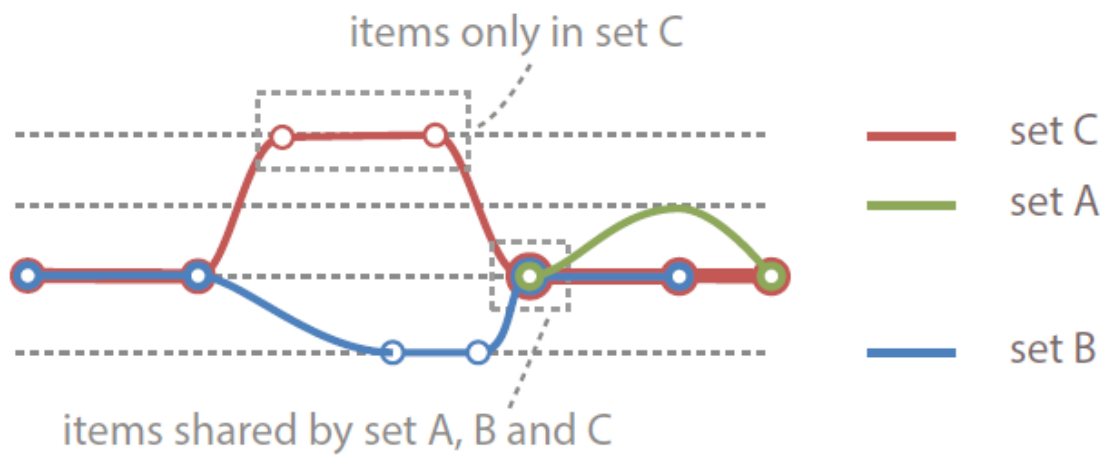


Fold small branches on MST

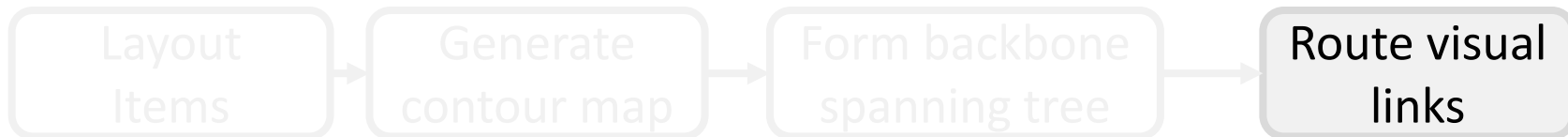


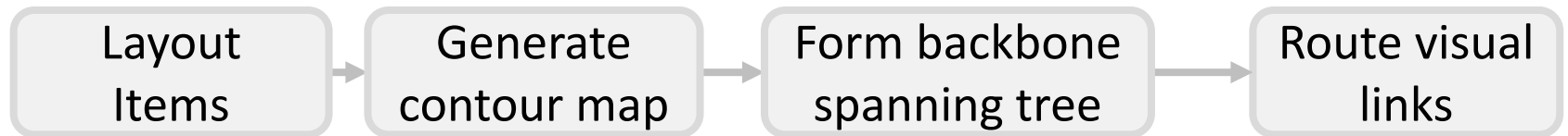
Straighten branches

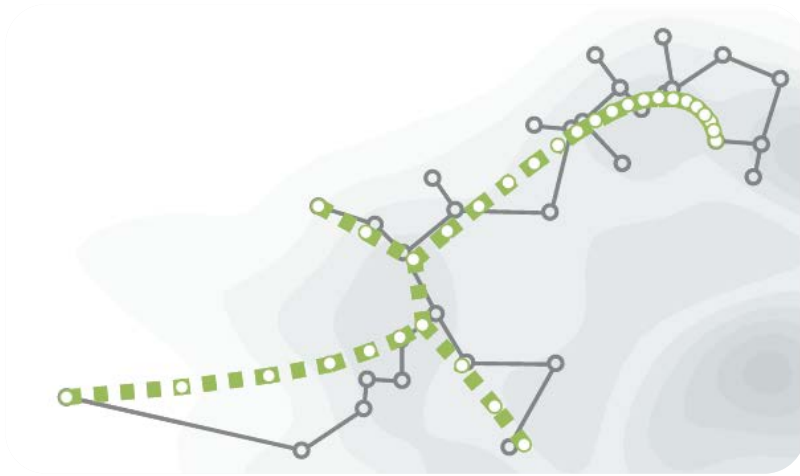




Draw visual link for individual sets



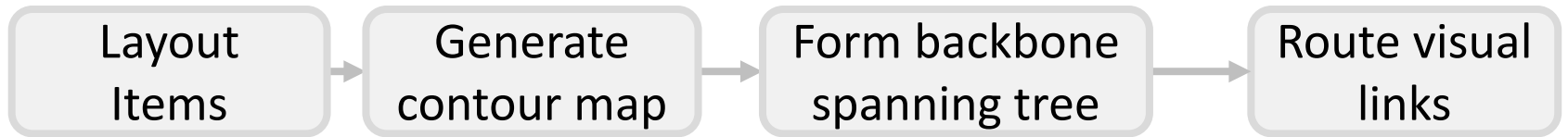




the original MST and the simplified backbone



the visual links for three sets





Motivation: data model and research questions

Approaches

Previous works

Technical details

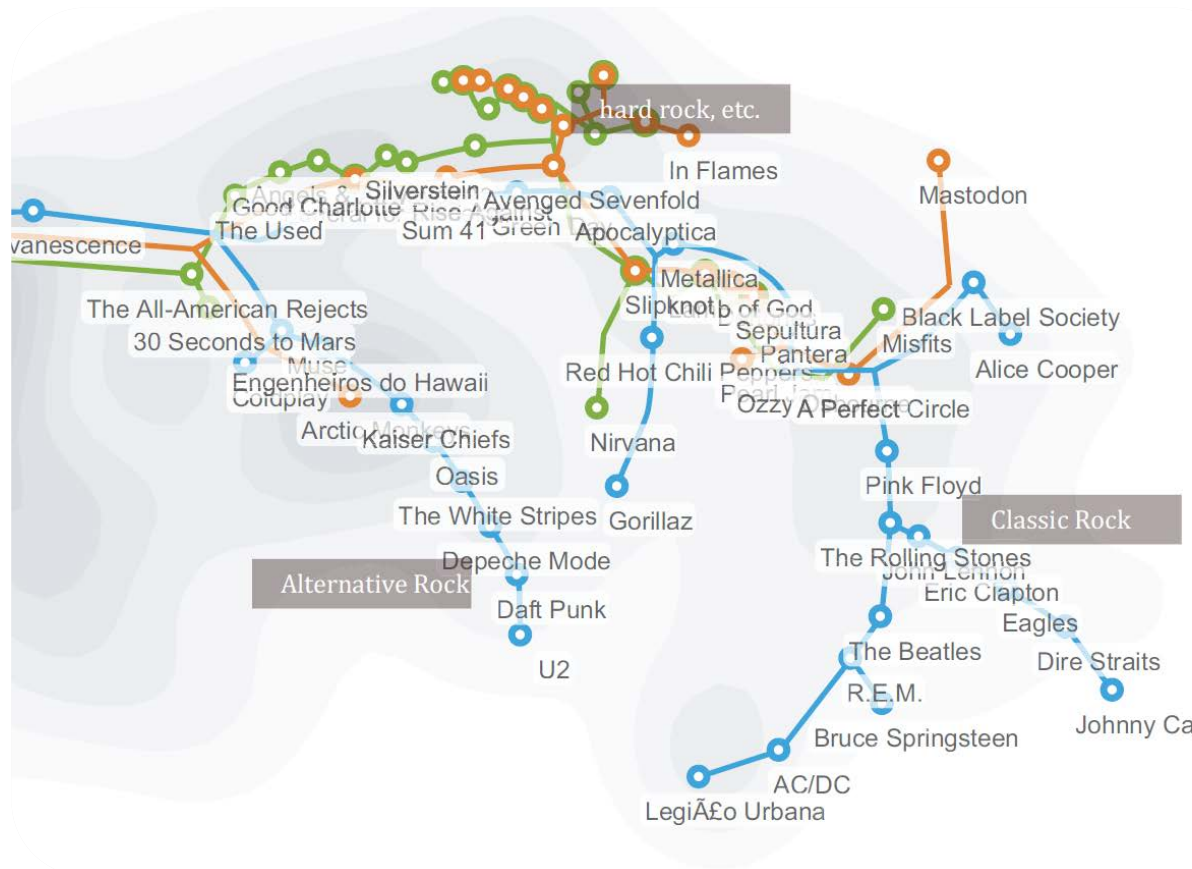
**Case studies**

Limitation and future works

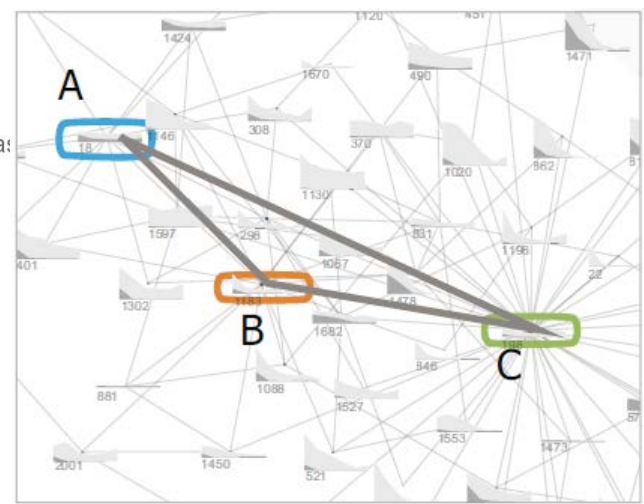
Outline



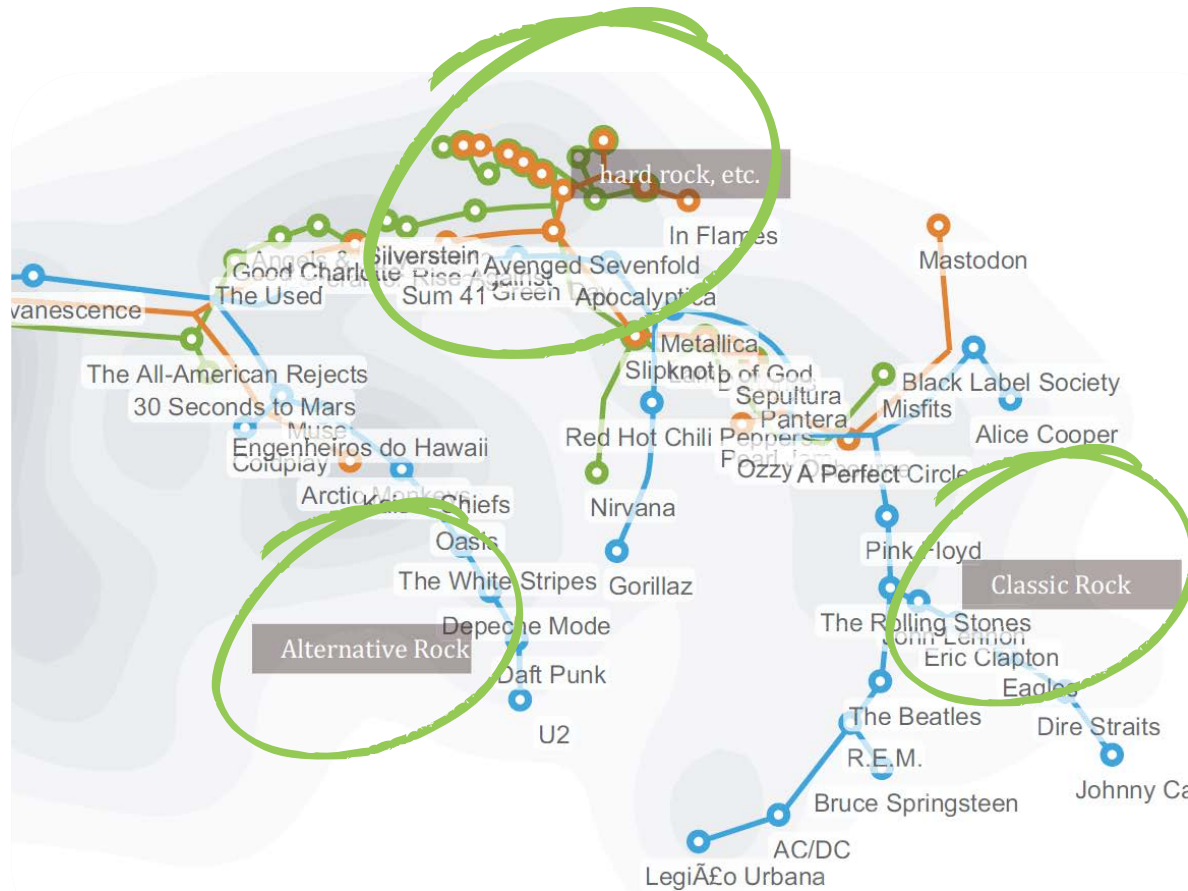




Last.fm  
 Artist similarity  
 User friendship  
 Listening history

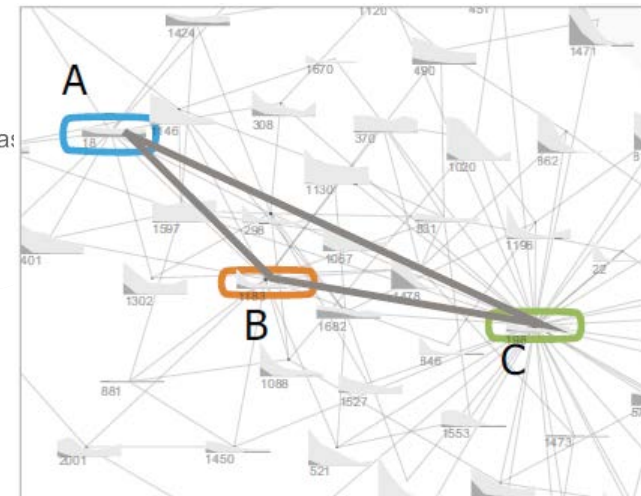


Social site data



Last.fm  
 Artist similarity  
 User friendship  
 Listening history

Social site data



**Glyph design** for homophily analysis

**Set visualization** over item clusters and layout algorithm

**Case studies**

Summary

Motivation: data model and research questions

Approaches

Previous works

Technical details

Case studies

**Limitation and future works**

Outline

# Scalability of glyph design

Use different graph layout, aggregate the nodes

Limitation & future works



## Scalability of glyph design

Use different graph layout, aggregate the nodes

## Scalability of set visualization

Improve layout algorithm

Limitation & future works

## Scalability of glyph design

Use different graph layout, aggregate the nodes

## Scalability of set visualization

Improve layout algorithm

## Evaluation

Compare with existing techniques (Line set, Kelp diagram)

Limitation & future works

## Scalability of glyph design

Use different graph layout, aggregate the nodes

## Scalability of set visualization

Improve layout algorithm

## Evaluation

Compare with existing techniques (Line set, Kelp diagram)

## Application of set visualization technique

Draw sets on word cloud, tree map, etc.

## Limitation & future works



## Last.fm Data

Artist similarity collected through Last.fm web API

User information could also be accessed

## Infovis 04 publication data

Keyword similarity: through topic modeling (LDA) and co-citation

Dataset collection & processing



# The Great Bear



# Metro map drawing