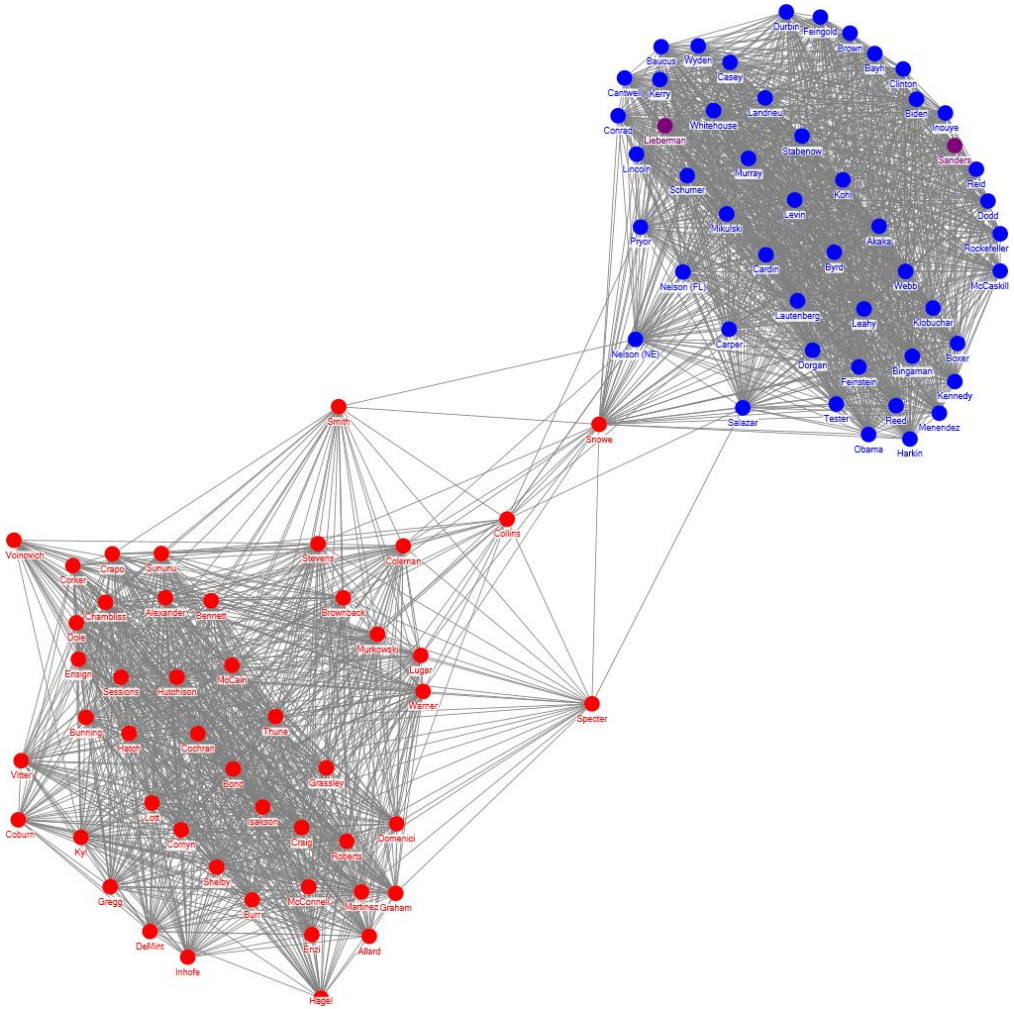
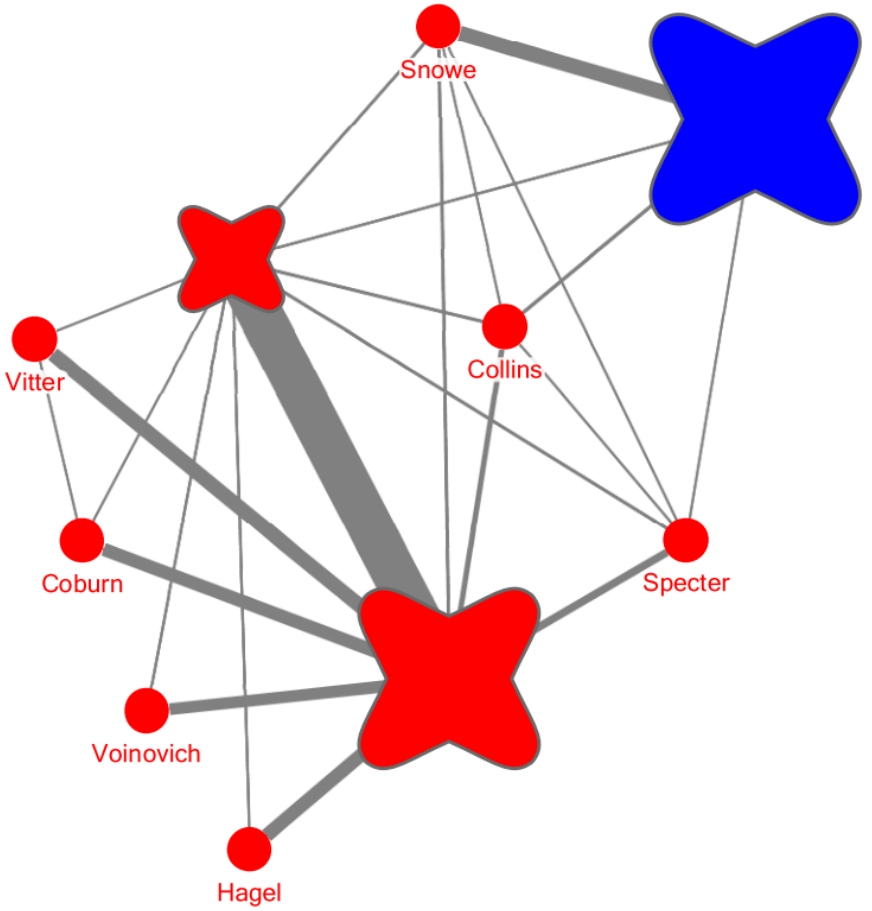


CS 7280-03 Special Topics on Visualization in Network Science

Lecture 10



Professor Cody Dunne

https://codydunne.github.io/cs7280-f16/

c.dunne@northeastern.edu

Homework 3

https://codydunne.github.io/cs7280-f16/hw/Homework-3-D3-spring-layout

Web Workers

Project Discussion

https://codydunne.github.io/cs7280-f16/project

Discussion: PivotGraph

Visual Exploration of Multivariate Graphs

Martin Wattenberg
Visual Communication Lab, IBM Research
1 Rogers St., Cambridge MA 02142
mwatten@us.ibm.com

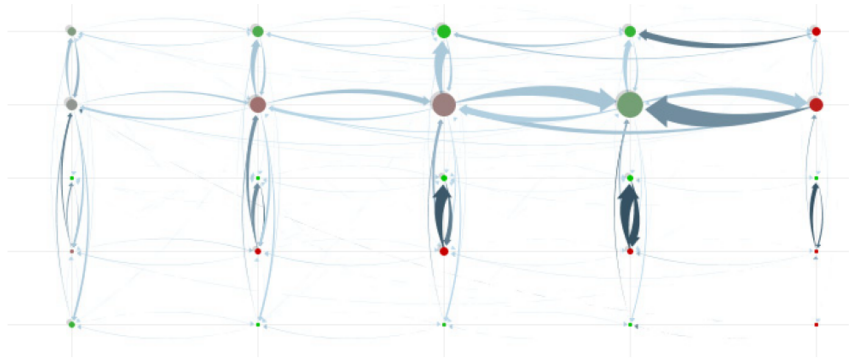


Figure 1. *A PivotGraph visualization of a large graph rolled up onto two categorical dimensions*

ABSTRACT
This paper introduces PivotGraph, a software tool that uses a new technique for visualizing and analyzing graph structures. The technique is designed specifically for graphs that are "multivariate," i.e., where each node is associated with several attributes. Unlike visualizations which emphasize global graph topology, PivotGraph uses a simple grid-based approach to focus on the relationship between node attributes and connections. The interaction technique is derived from an analogy with methods seen in spreadsheet pivot tables and in online analytical processing (OLAP). Finally, several examples are presented in which PivotGraph was applied to real-world data sets.

Author Keywords
information visualization, graph drawing

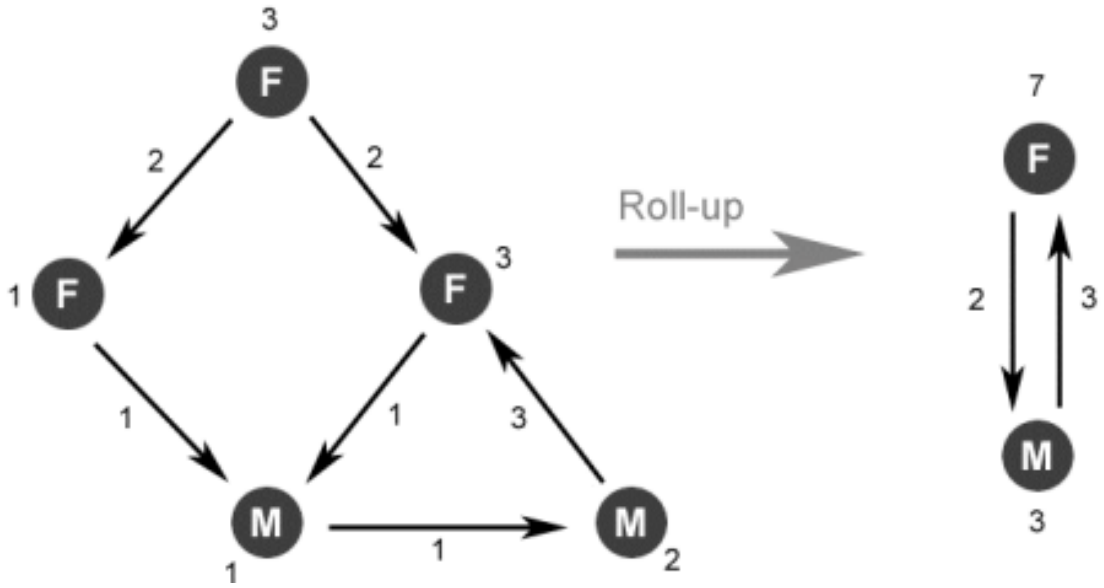
ACM Classification Keywords

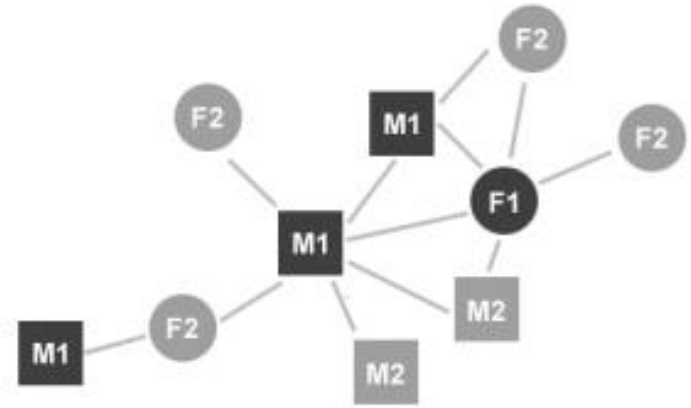
H.5.2. [Information Interfaces]: User Interfaces. I.3.6 [Methodology and Techniques]: Interaction Techniques

INTRODUCTION
A common goal in exploratory data analysis is to form hypotheses about a graph. In contexts ranging from social networks to Markov chains, it can be important to understand the properties of a graph and make inferences about why it looks the way it does. Because of the complex structure and large size of many graphs, such an analysis can be difficult.

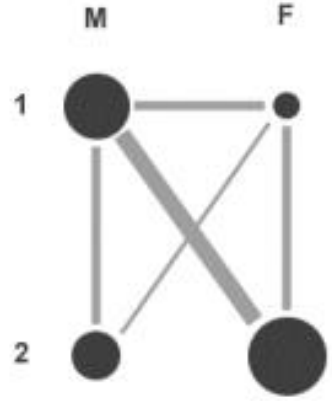
Visualization is one solution to this problem. In one popular approach, nodes of a graph are represented by a dot, and edges by lines or curves linking them. Such "node-and-link diagrams" have been explored intensively throughout this century (see [3, 10, 14, 15]). A second common approach is a matrix view, or density table, that displays a graph's adjacency matrix overlaid on a grid [4].

Both node-and-link diagrams and matrix views emphasize link structure. In many cases, however, data about a graph





Node and Link Diagram



PivotGraph Roll-up

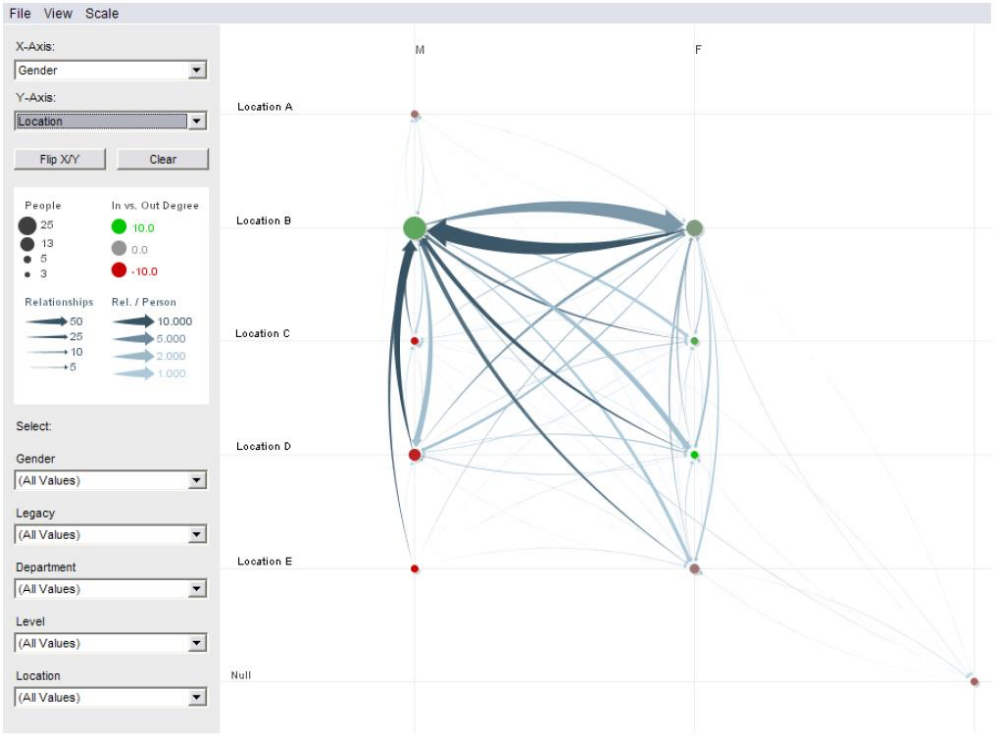


Figure 6. *A one-dimensional layout*

[Wattenberg, 2006]

Discussion: GLO-STIX

GLO-STIX: Graph-Level Operations for Specifying Techniques and Interactive eXploration

Charles D. Stolper, Minsuk Kahng, Zhiyuan Lin, Florian Foerster, Aakash Goel, John Stasko, and Duen Horng Chau

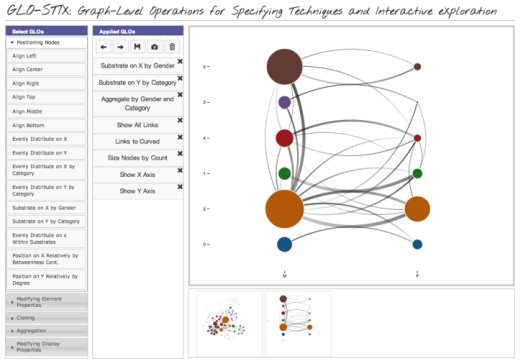


Fig. 1: A screenshot of the GLO-STIX user interface showing a user exploring the Les Misérables character co-occurrence graph using graph-level operations (GLOs). Nodes are characters, and an edge connects two characters if they co-occur in a chapter. The original node-link view of the graph is saved by the user as a snapshot in the bottom pane. From the list of operations available (shown in left-most column), applying those selected in the middle column transforms the original graph into the PivotGraph visualization [30] displayed in the main view. All graph figures in this paper were generated using GLO-STIX.

Abstract— The field of graph visualization has produced a wealth of visualization techniques for accomplishing a variety of analysis tasks. Therefore analysts often rely on a suite of different techniques, and visual graph analysis application builders strive to provide this breadth of techniques. To provide a holistic model for specifying network visualization techniques (as opposed to considering each technique in isolation) we present the *Graph-Level Operations* (*GLO*) model. We describe a method for identifying GLOs and apply it to identify five classes of GLOs, which can be flexibly combined to re-create six canonical graph visualization techniques. We discuss advantages of the GLO model, including potentially discovering new, effective network visualization techniques and easing the engineering challenges of building multi-technique graph visualization applications. Finally, we implement the GLOs that we identified into the GLO-STIX prototype system that enables an analyst to interactively explore a graph by applying GLOs.

Index Terms—Graph-level operations, graph visualization, visualization technique specification, graph analysis, information visualization.

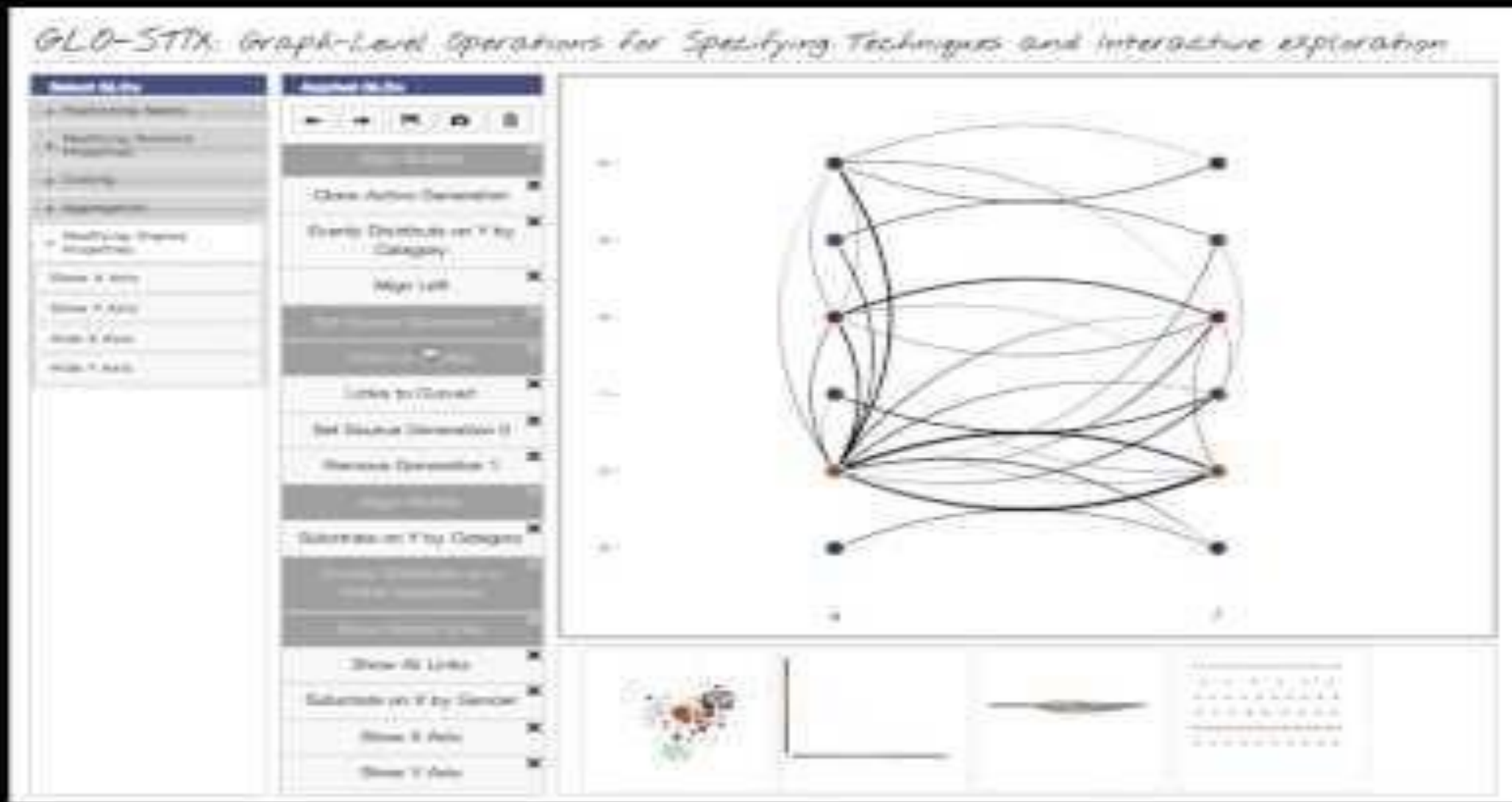
- *Charles D. Stolper is with the College of Computing, Georgia Institute of Technology. E-mail: chadstolper@gatech.edu.*
- *Minsuk Kahng is with the College of Computing, Georgia Institute of Technology. E-mail: kahng@gatech.edu.*
- *Zhiyuan Lin is with the College of Computing, Georgia Institute of Technology. E-mail: zlin48@gatech.edu.*
- *Florian Foerster is with the College of Computing, Georgia Institute of Technology. E-mail: florian.foerster@gatech.edu.*
- *Aakash Goel is with the College of Computing, Georgia Institute of Technology. E-mail: aakashgoel@gatech.edu.*
- *John Stasko is with the the College of Computing, Georgia Institute of Technology. E-mail: stasko@cc.gatech.edu.*
- *Duen Horng Chau is with the College of Computing, Georgia Institute of Technology. E-mail: polo@gatech.edu.*

Manuscript received 31 Mar. 2014; accepted 1 Aug. 2014; date of publication xx xxx 2014; date of current version xx xxx 2014.
For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.

1 Introduction

The field of graph visualization has provided analysts with a number of useful techniques for displaying the nodes and edges of a graph. Each of these techniques can be quite effective at showing aspects of the graph to the analyst. In other words, different techniques are effective at accomplishing different tasks. When analysts wish to perform multiple tasks, they often turn to multiple graph visualization techniques. Developers of graph visualization systems, in turn, must implement this variety of techniques in their applications.

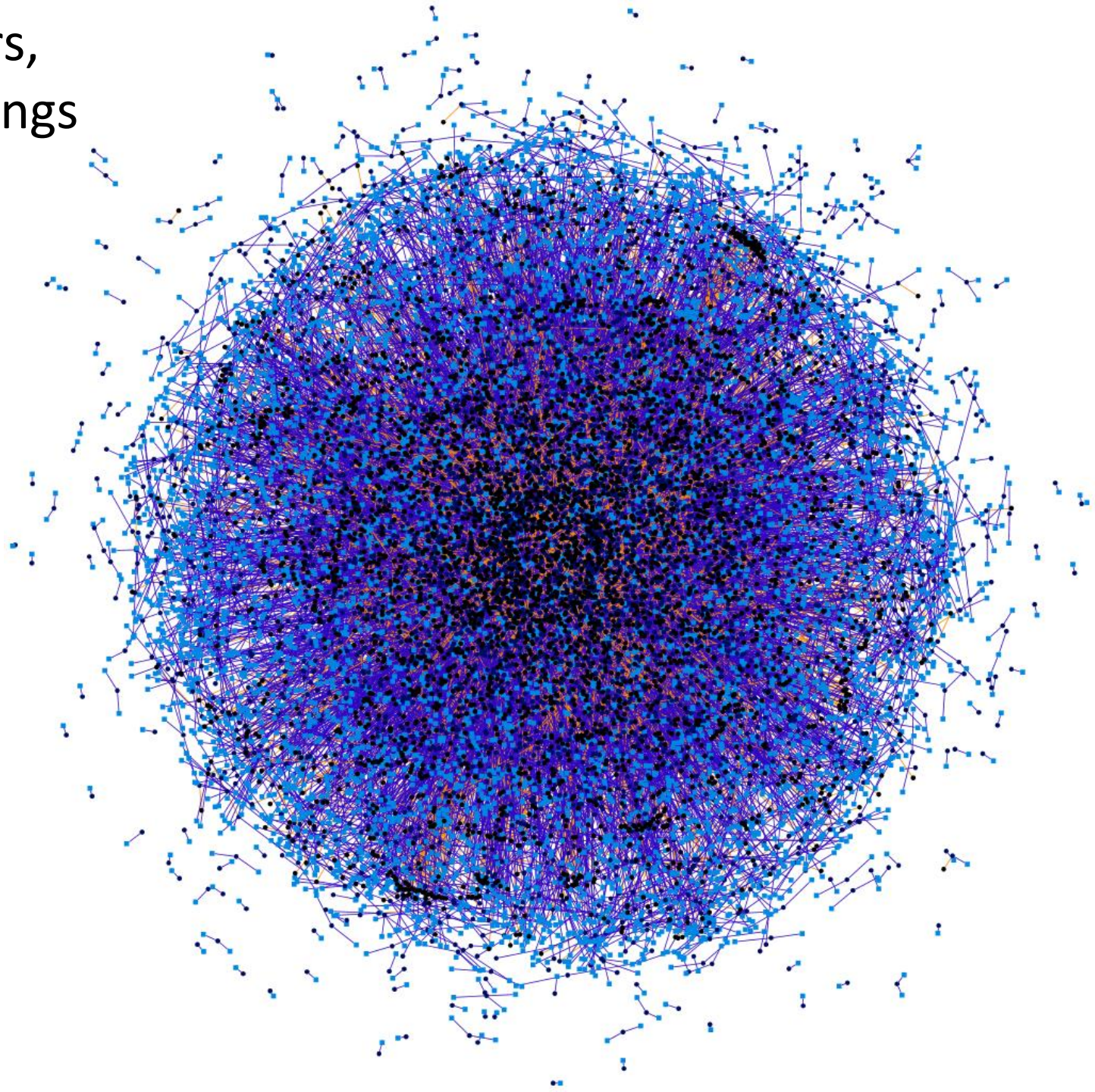
We introduce the new idea of **graph-level operations** (GLOs), which provides an alternative to implementing each graph visualization technique in isolation. GLOs are encapsulated manipulations of a graph visualization. Let us consider an example GLO: *positioning each node's glyph relatively on an axis according to a continuous attribute of the node*. This GLO might be used to stratify the nodes according to their node types, as in a *semantic substrates* visualization [22], which places nodes in non-overlapping regions, one region for each node type, to help reduce visual complexity (see Table 1, fourth visualization). The same GLO may also be used to position



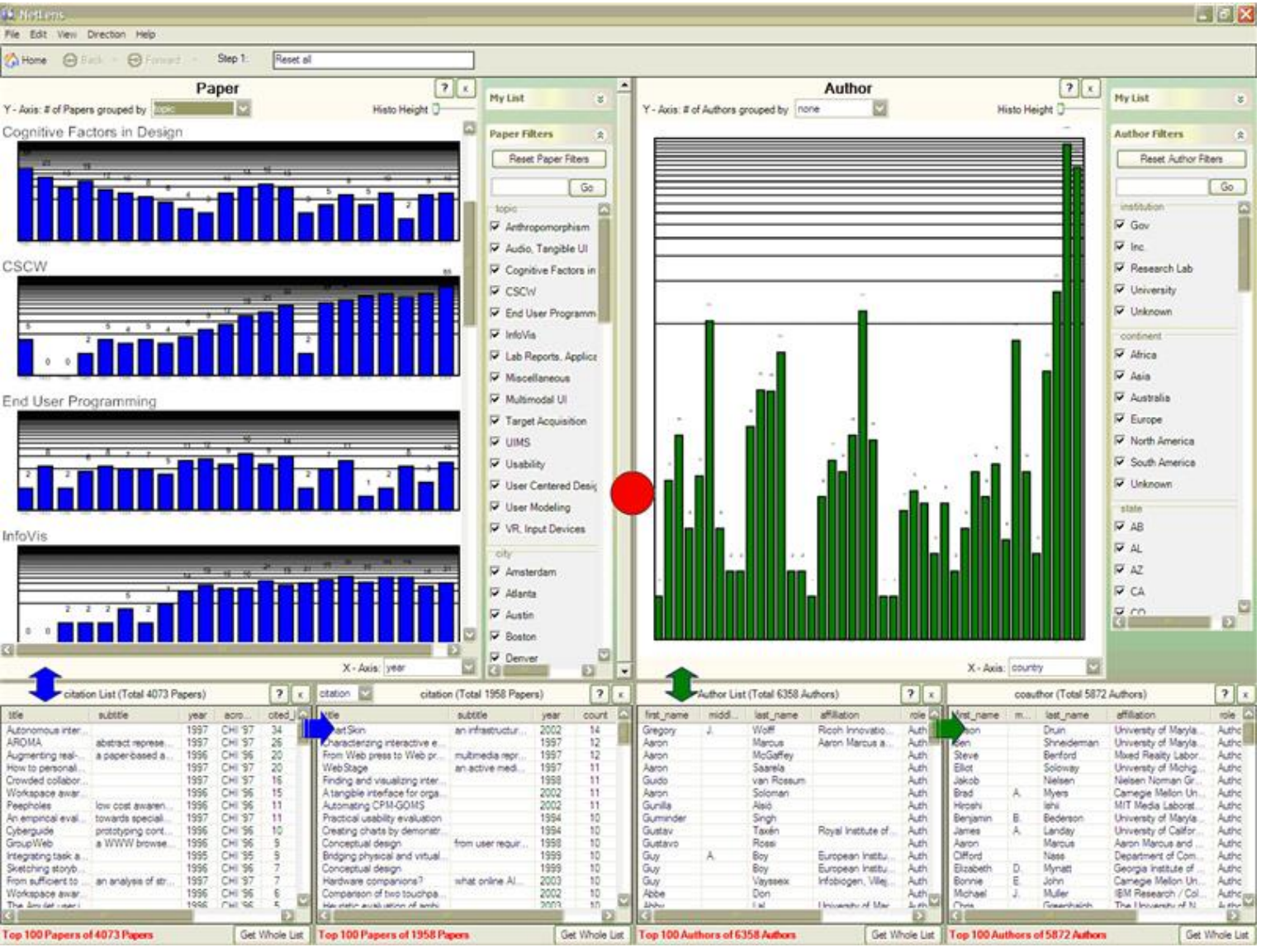
Stopler et al., 2014

Attribute Aggregation

ACM SigChi papers, authors, proceedings until 2004



NetLens



Video

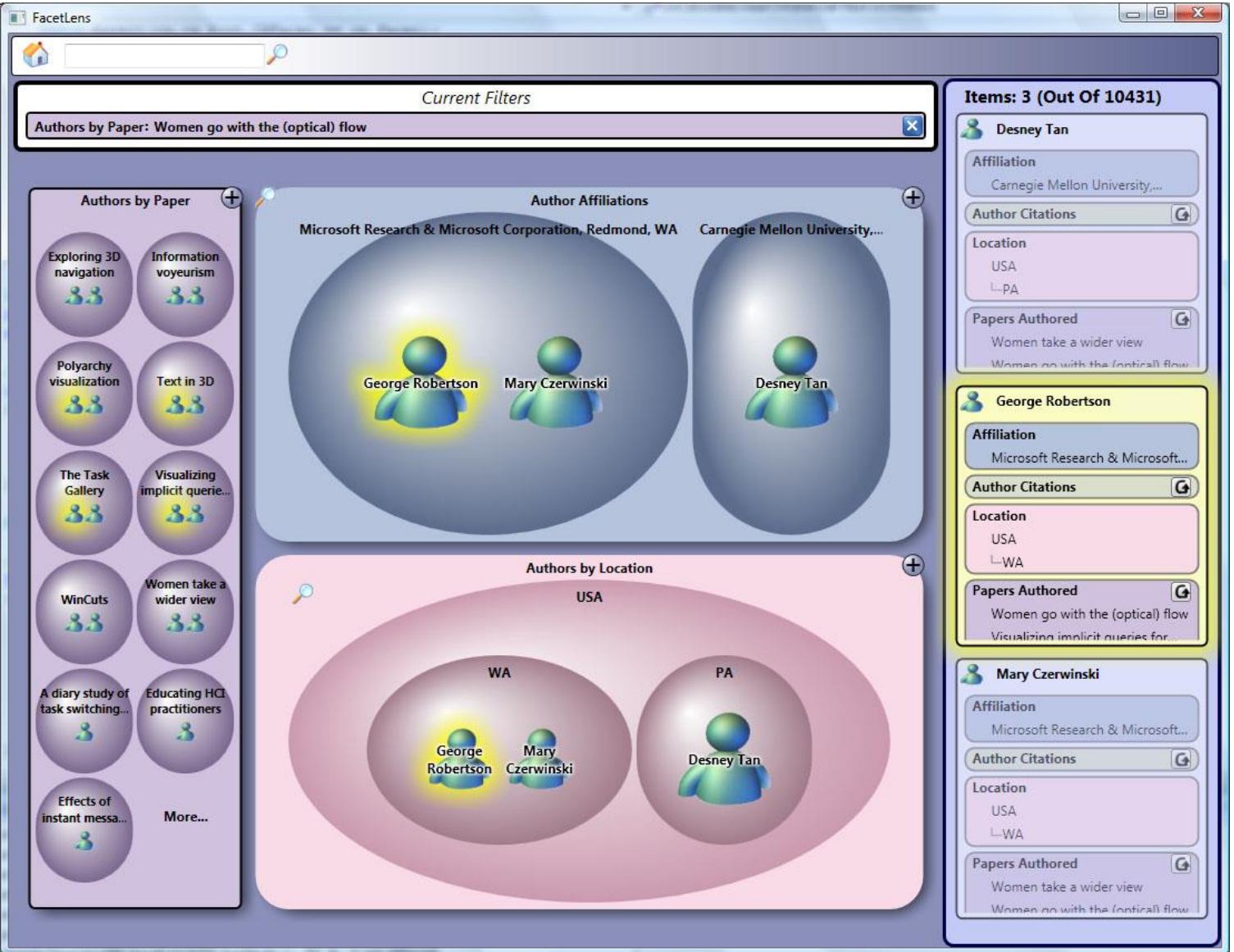
Kang et al., 2006

PaperLens

[Video]

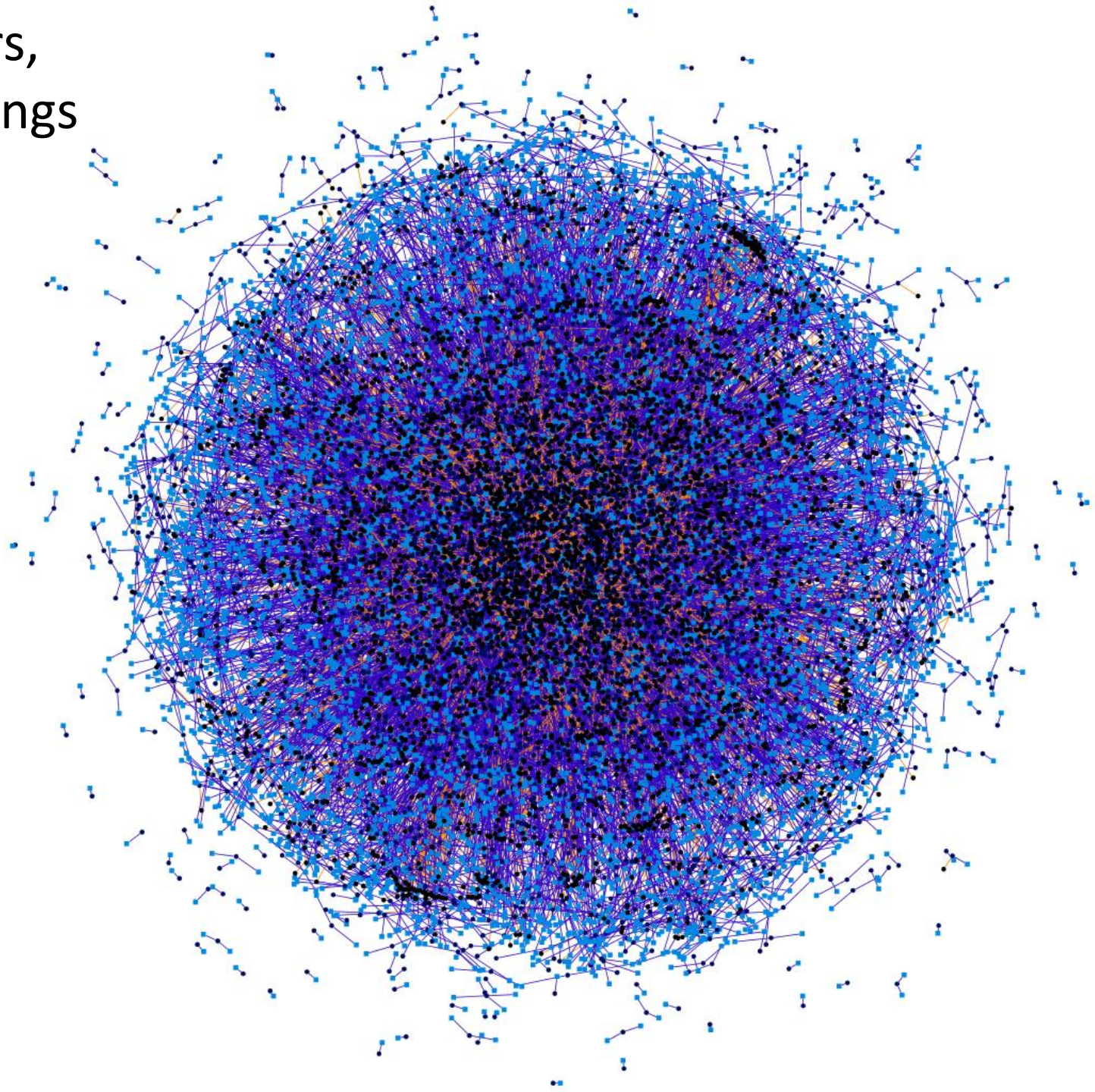
[Lee et al., 2004]

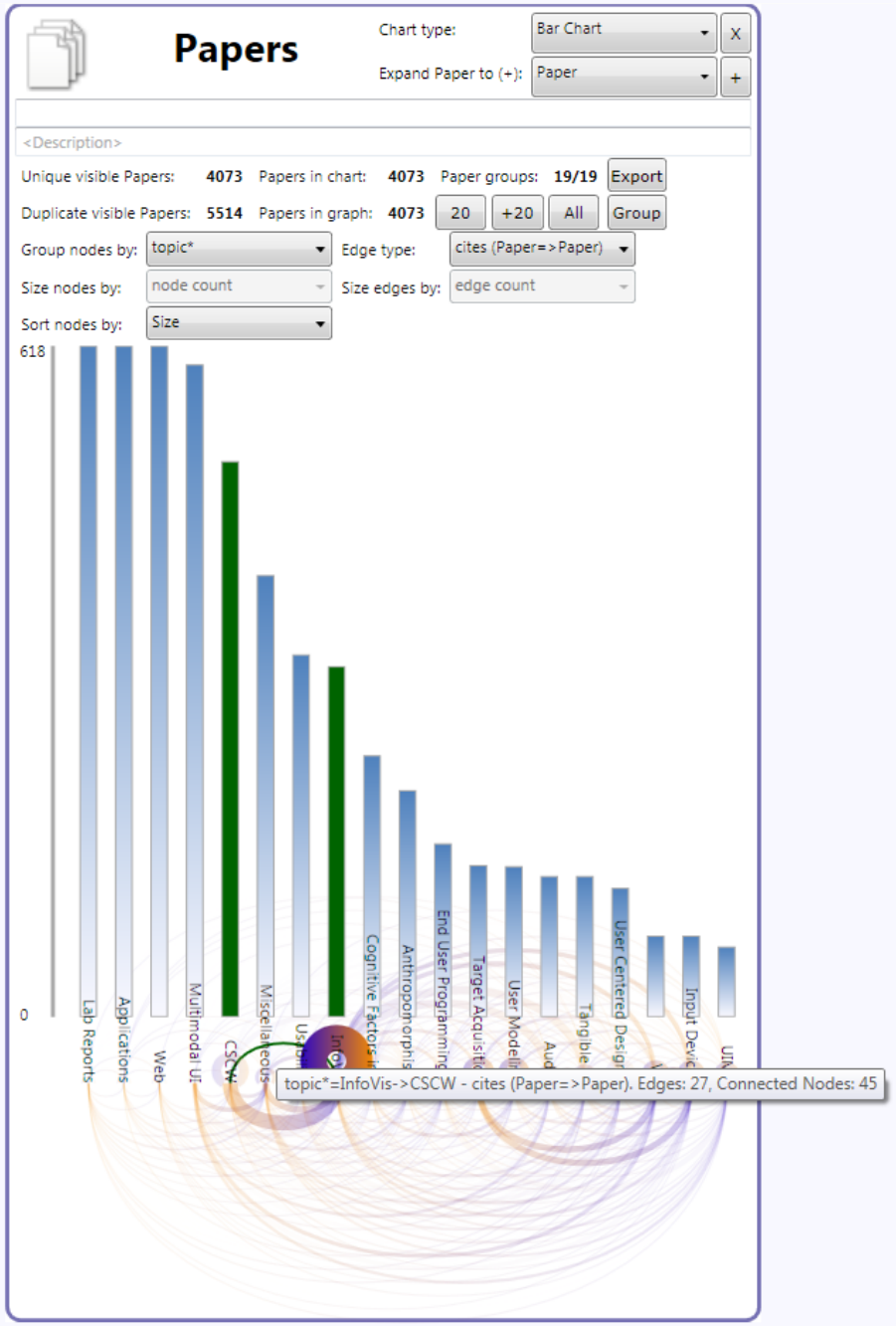
FacetLens

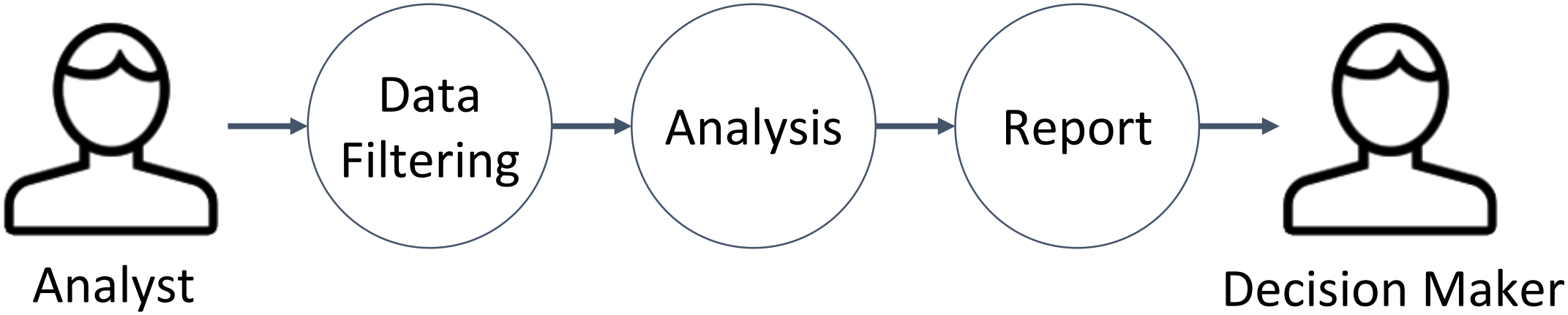


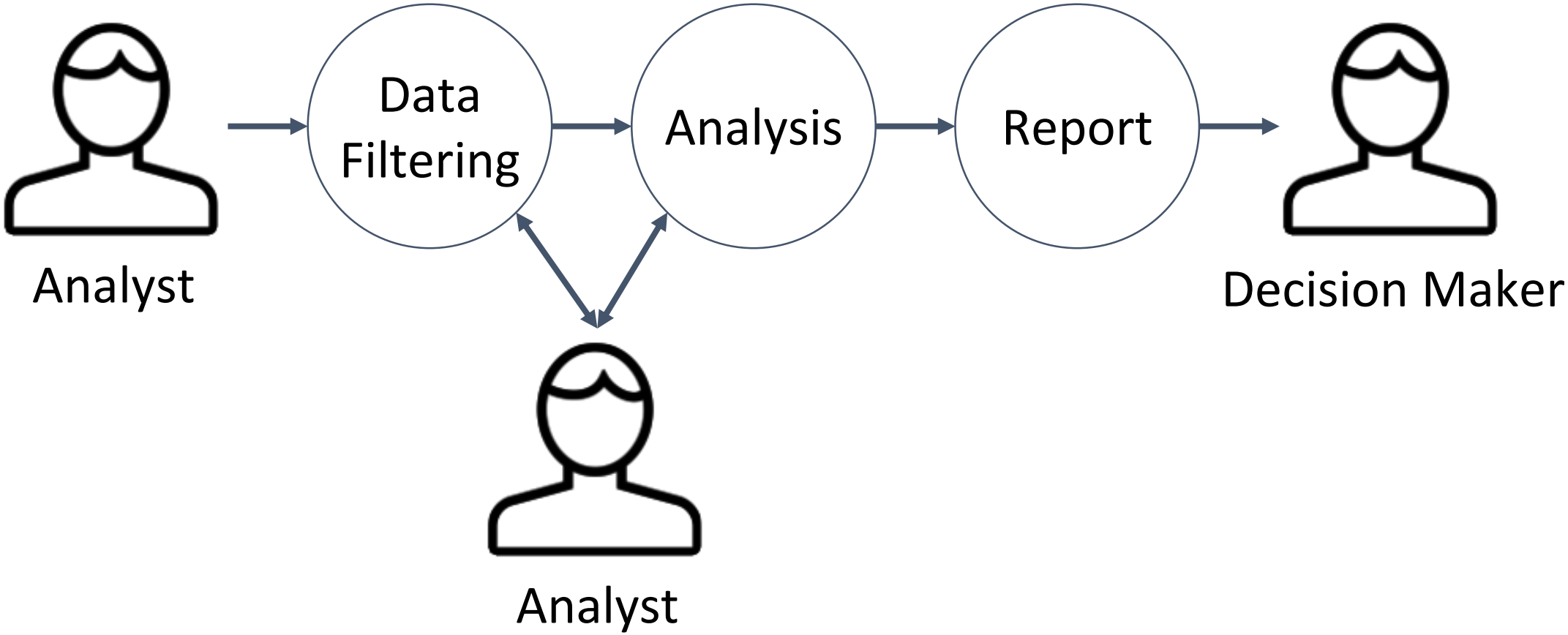
[Lee et al., 2009]

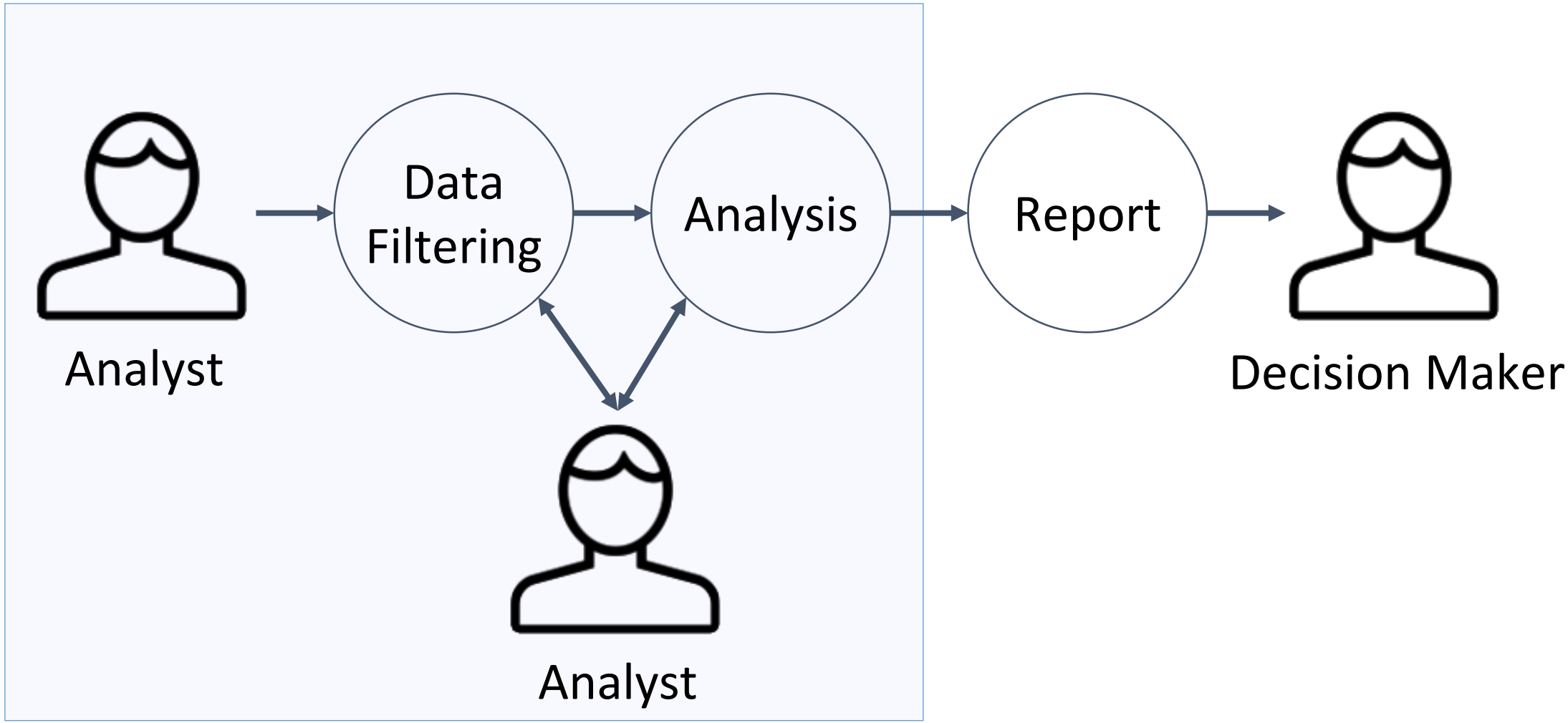
ACM SigChi papers, authors, proceedings until 2004

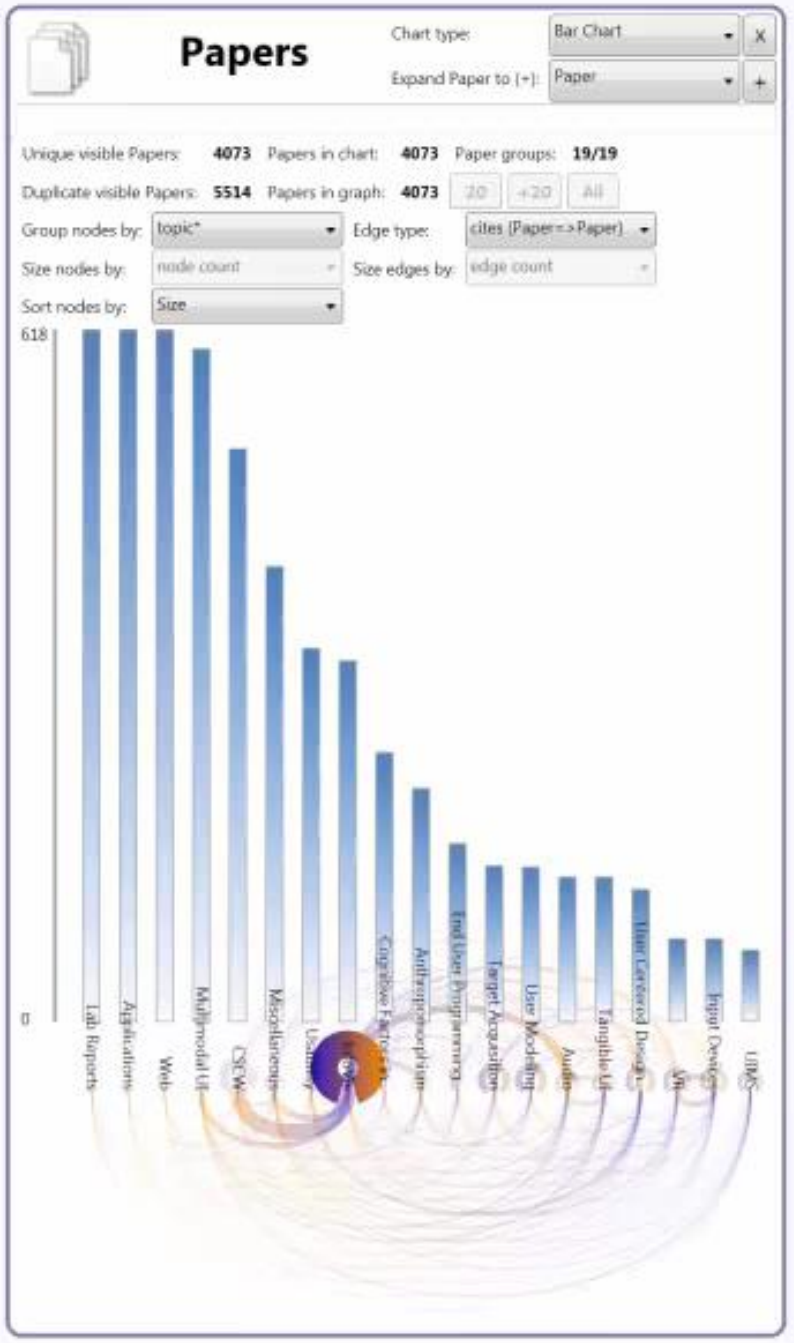



Papers
Chart type:
Bar Chart
X
Expand Paper to (+):
Paper
+
<Description>
Unique visible Papers: 4073
Papers in chart: 4073
Paper groups: 19/19
Export
Duplicate visible Papers: 5514
Papers in graph: 4073
20
+20
All
Group
Group nodes by: topic*
Edge type: cites (Paper=>Paper)
Size nodes by: node count
Size edges by: edge count
Sort nodes by: Size
618
0
Lab Reports
Applications
Web
Multimodal UI
CSCW
Miscellaneous
Cognitive Factors
Anthropomorphism
End User Programming
Target Acquisition
User Modeling
Tangible
User Centered Design
Input Devices
topic*=InfoVis->CSCW - cites (Paper=>Paper). Edges: 27, Connected Nodes: 45


Data
Filtering
Analysis
Report
Analyst
Decision Maker

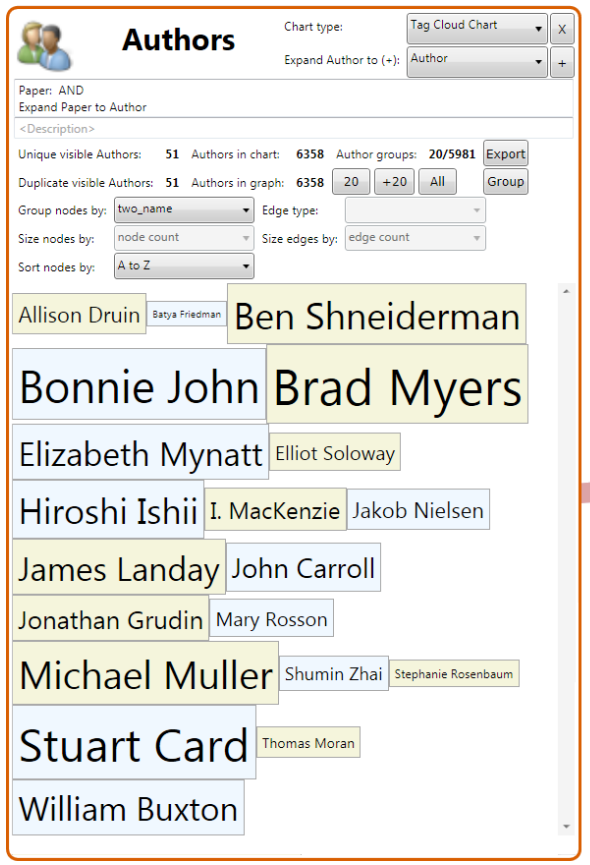

Data
Filtering
Analysis
Report
Analyst
Decision Maker
Analyst

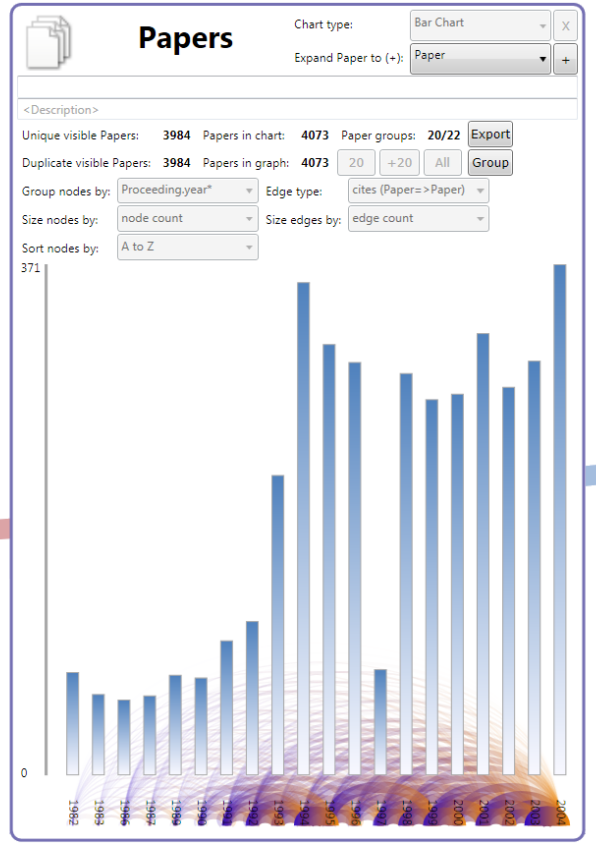

Data
Filtering
Analysis
Report
Analyst
Analyst
Decision Maker

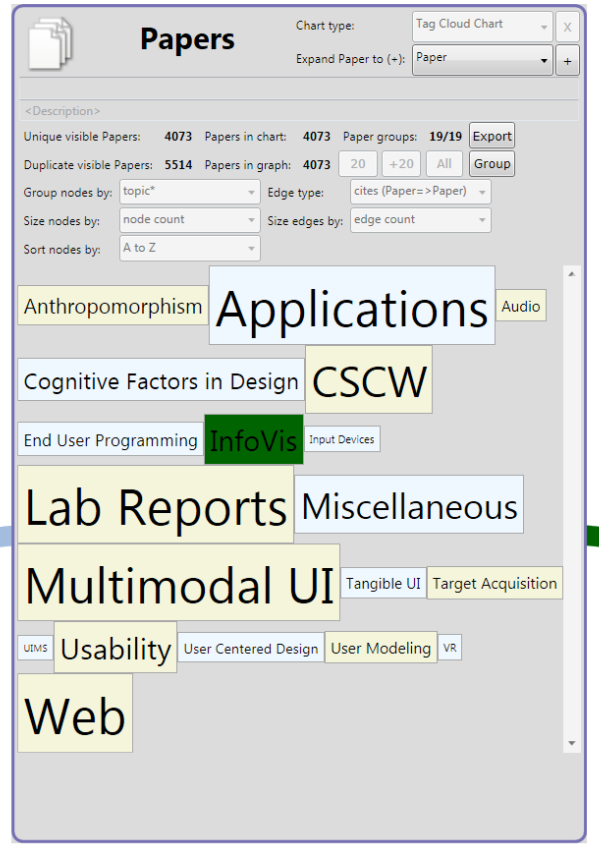

Papers
Chart type:
Bar Chart
Expand Paper to (+):
Paper
Unique visible Papers: 4073
Papers in chart: 4073
Paper groups: 19/19
Duplicate visible Papers: 5514
Papers in graph: 4073
20
+20
All
Group nodes by: topic*
Edge type: cites (Paper=>Paper)
Size nodes by: node count
Size edges by: edge count
Sort nodes by: Size
618
0
Lab Reports
Applications
Web
Multimodal UI
CSCW
Miscellaneous
Usability
Cognitive Factors
Anthropomorphism
End User Programming
Target Acquisition
User Modeling
Audio
Tangible UI
User Centered Design
VR
Input Devices
UIMS

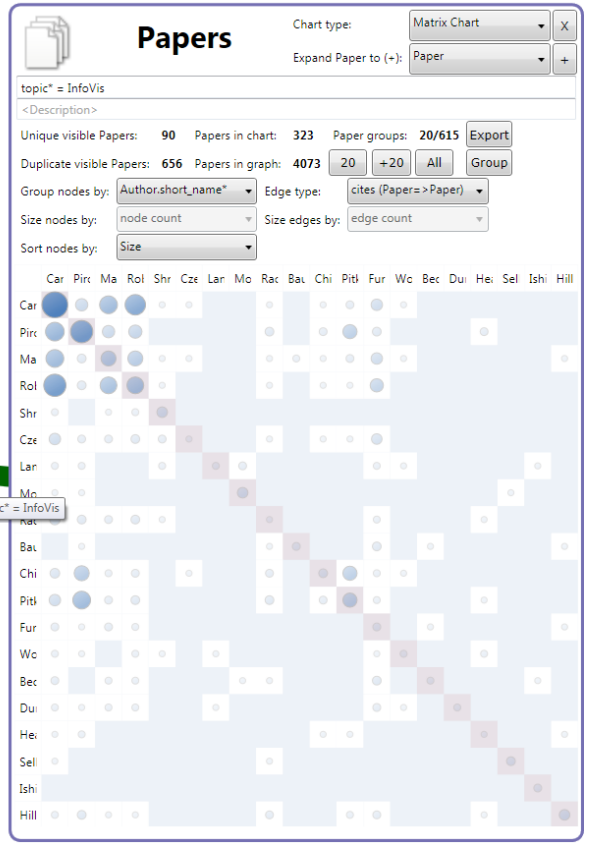
GraphTrail

Aggregating charts



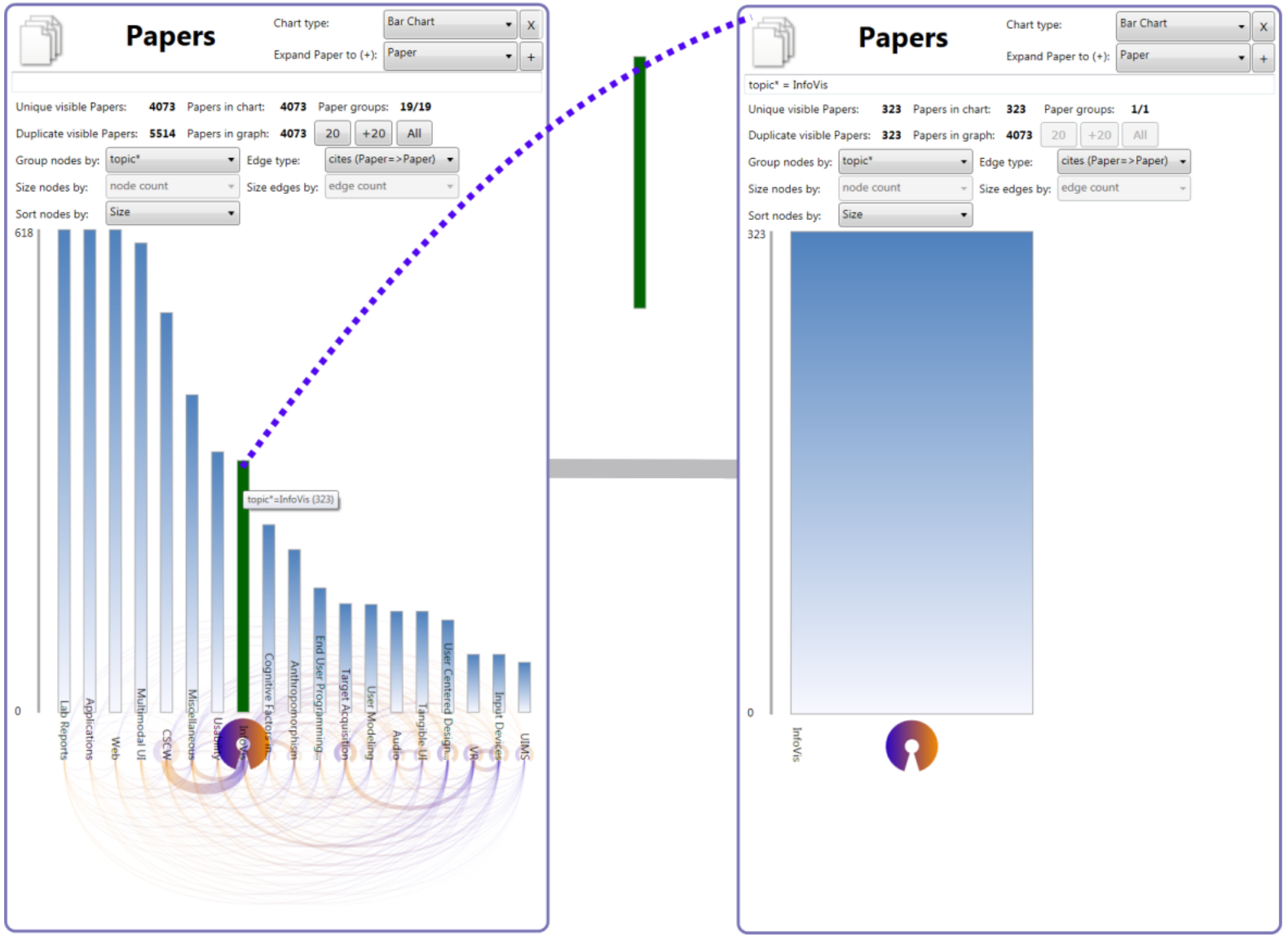






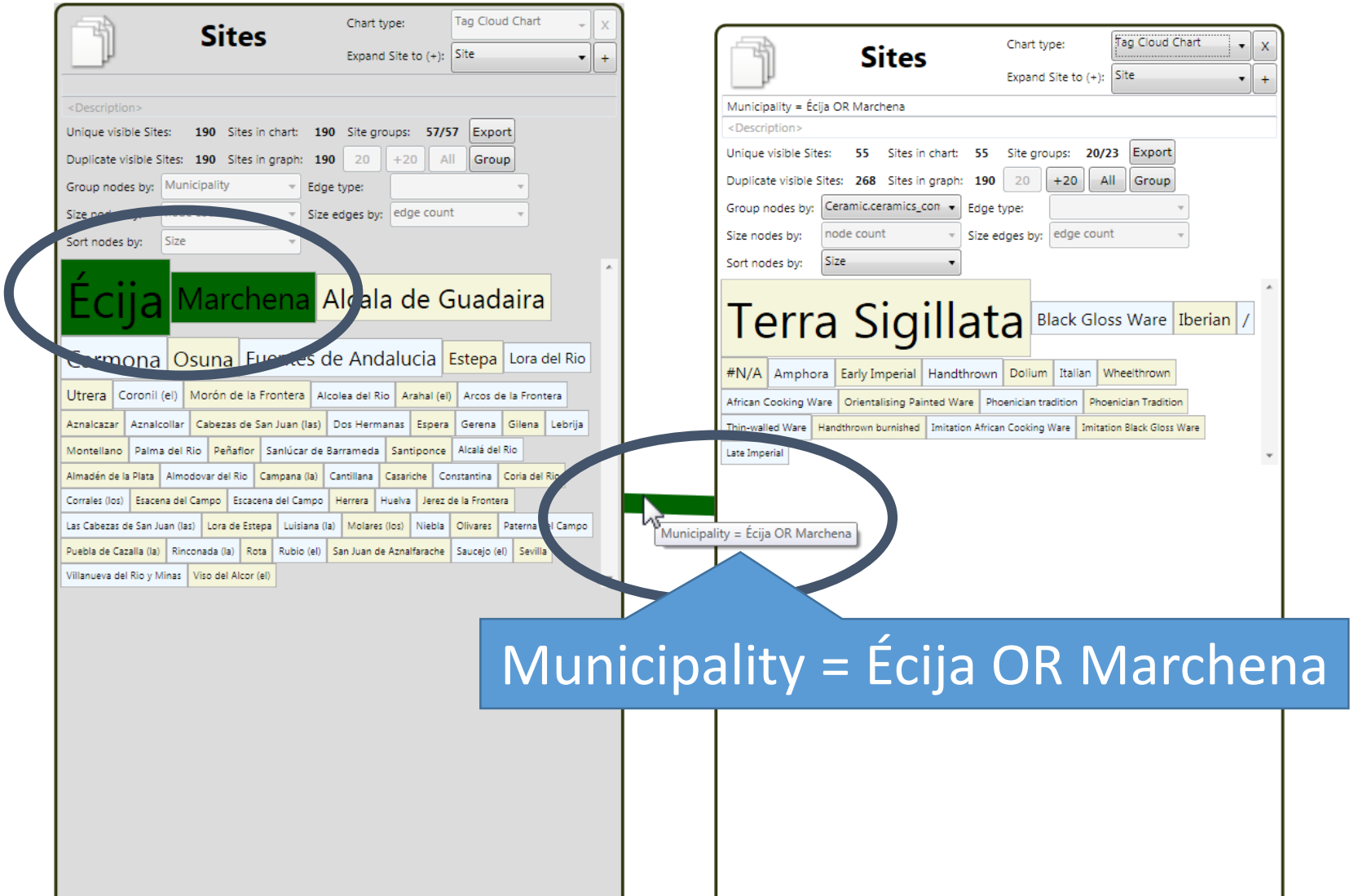
GraphTrail

Drag and drop interaction



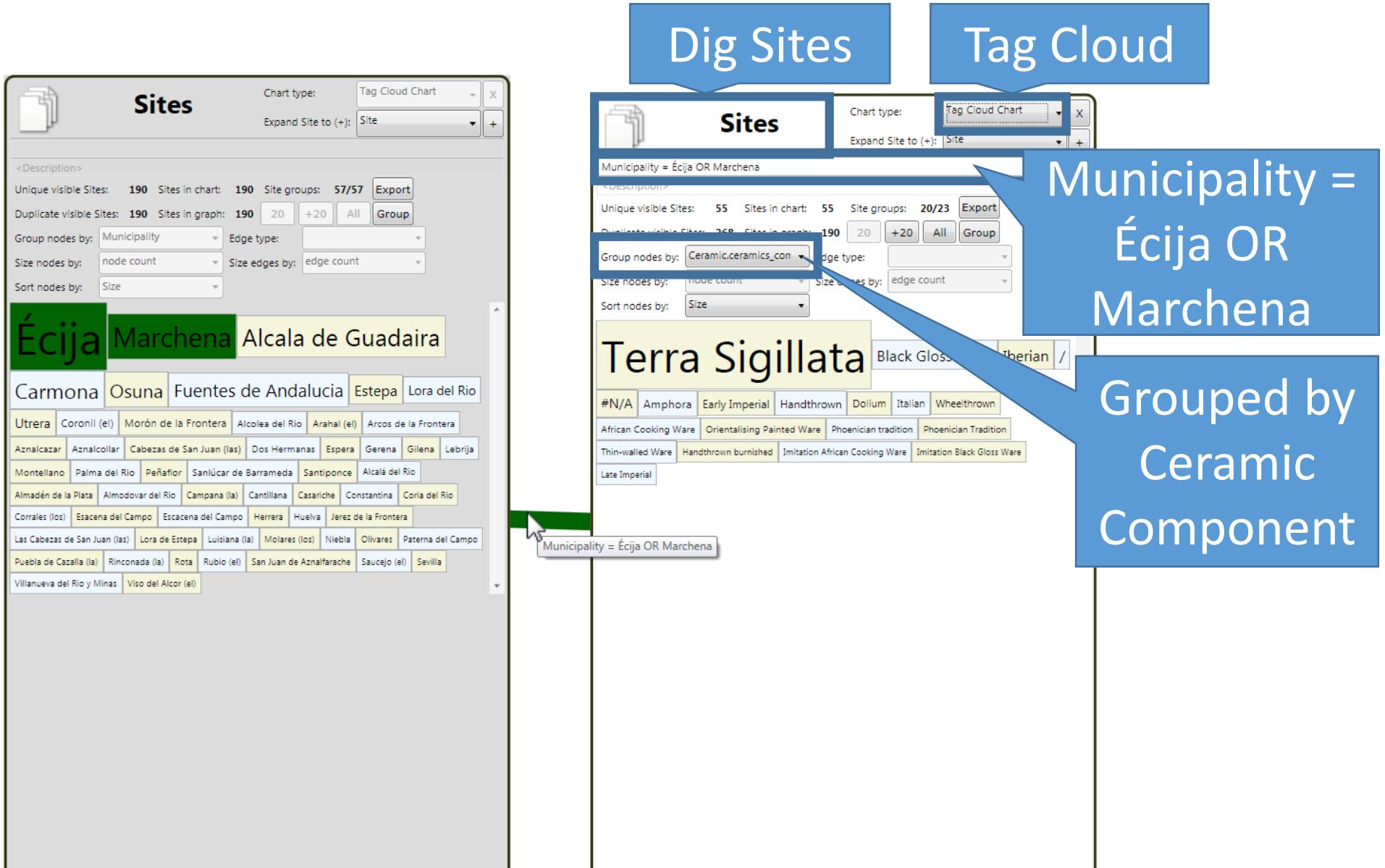
GraphTrail

Provenance/history visualization



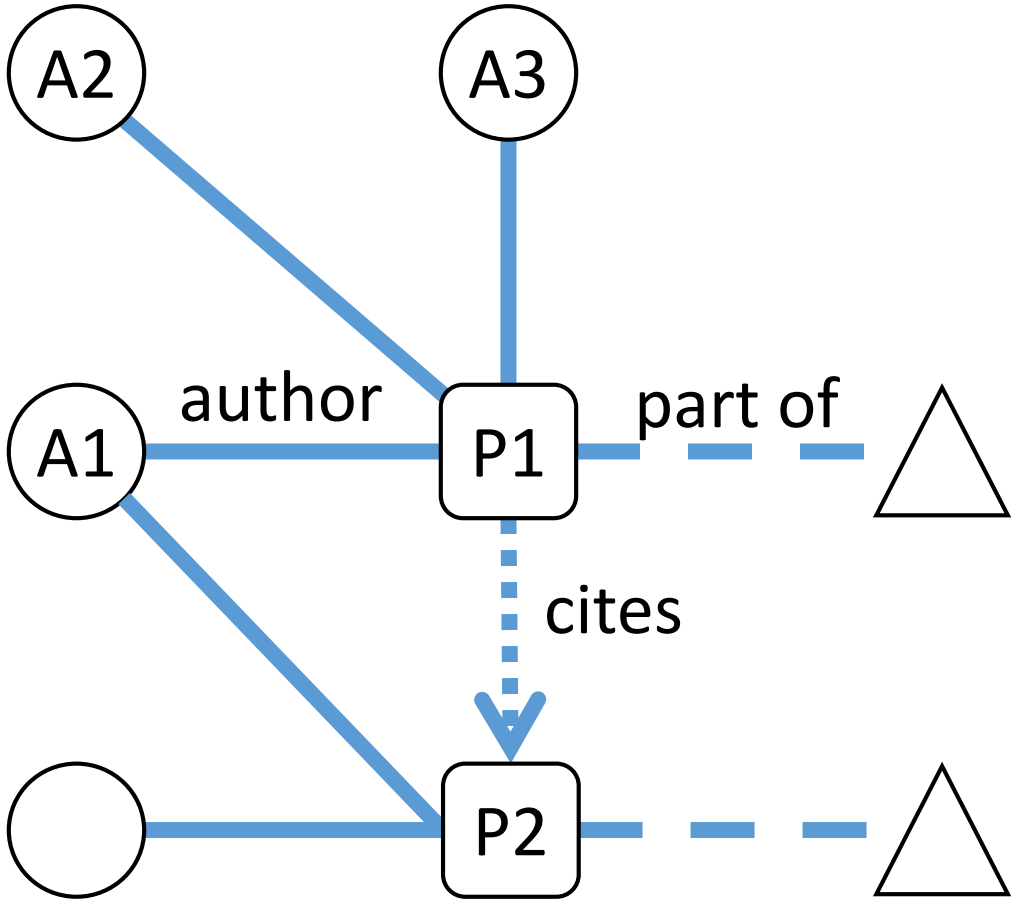
GraphTrail

Provenance & chart parameterization



GraphTrail

Pivoting & Derived Attributes

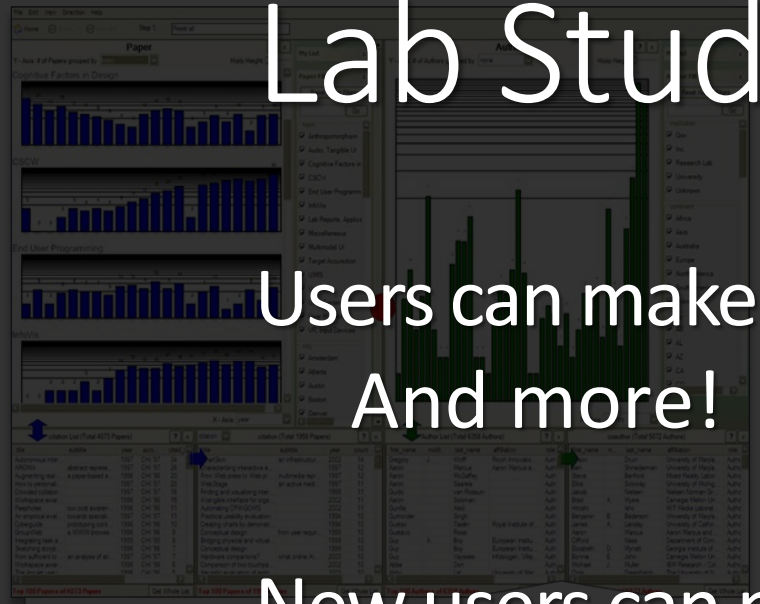


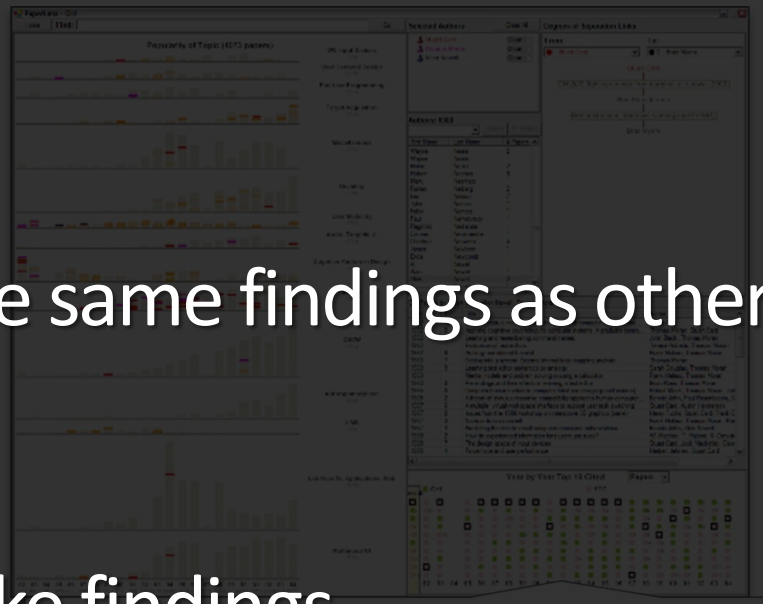
Author
- Name
- Affiliation
- Paper.Topic*

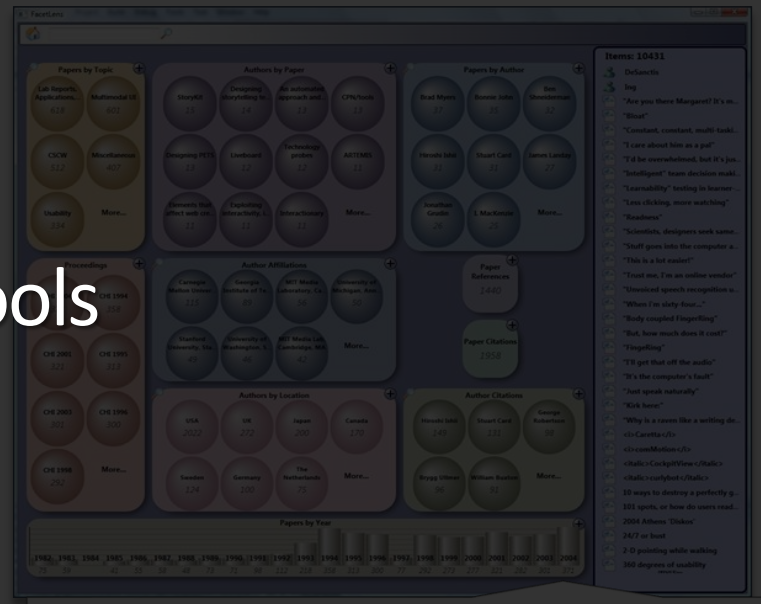
Paper
- Topic
- Year

Proceeding
- Country
- Date

Lab Study







Users can make the same findings as other tools

And more!

New users can make findings

NetLens

PaperLens

FacetLens

New users understand the exploration history

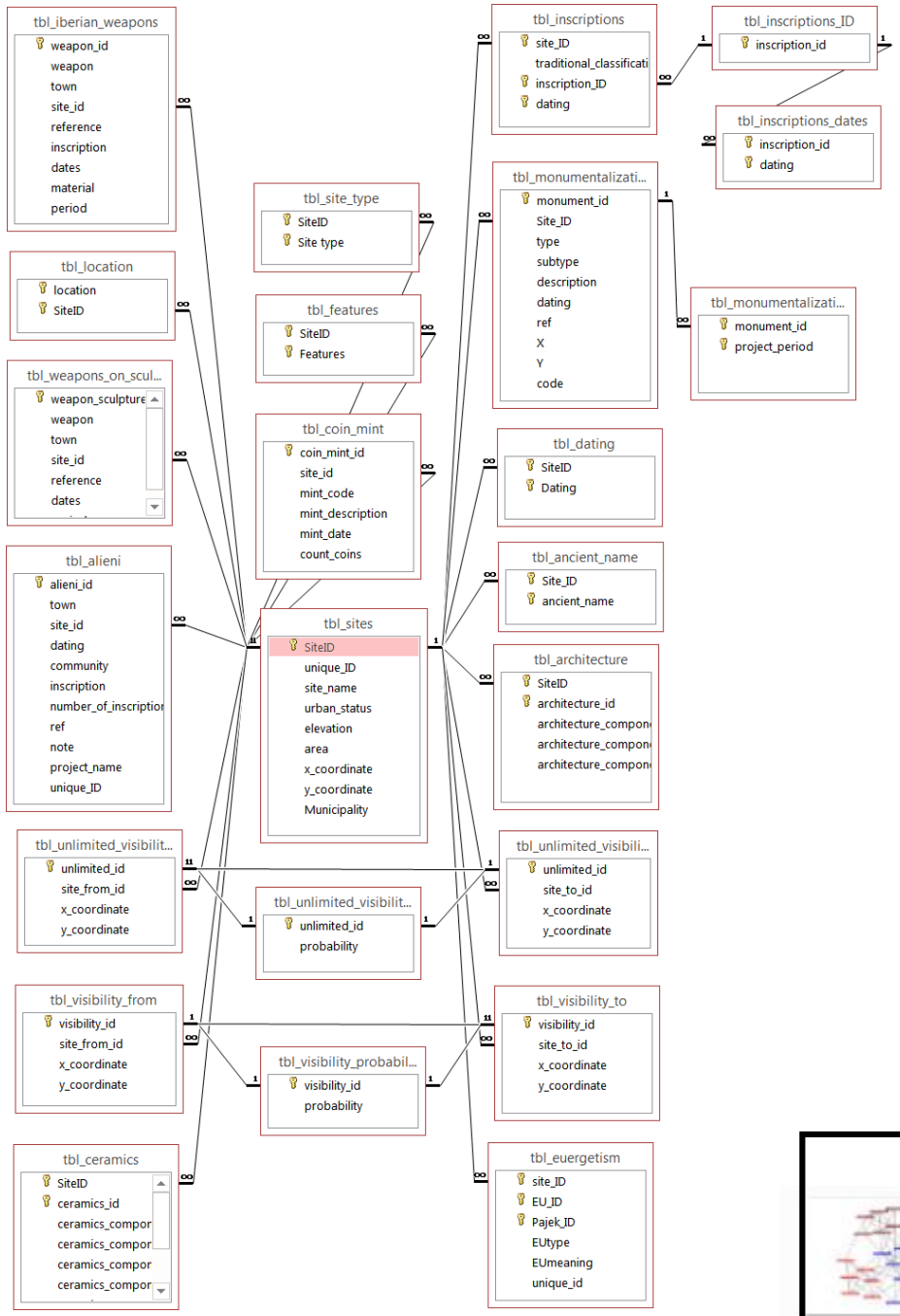
And usually motivation!

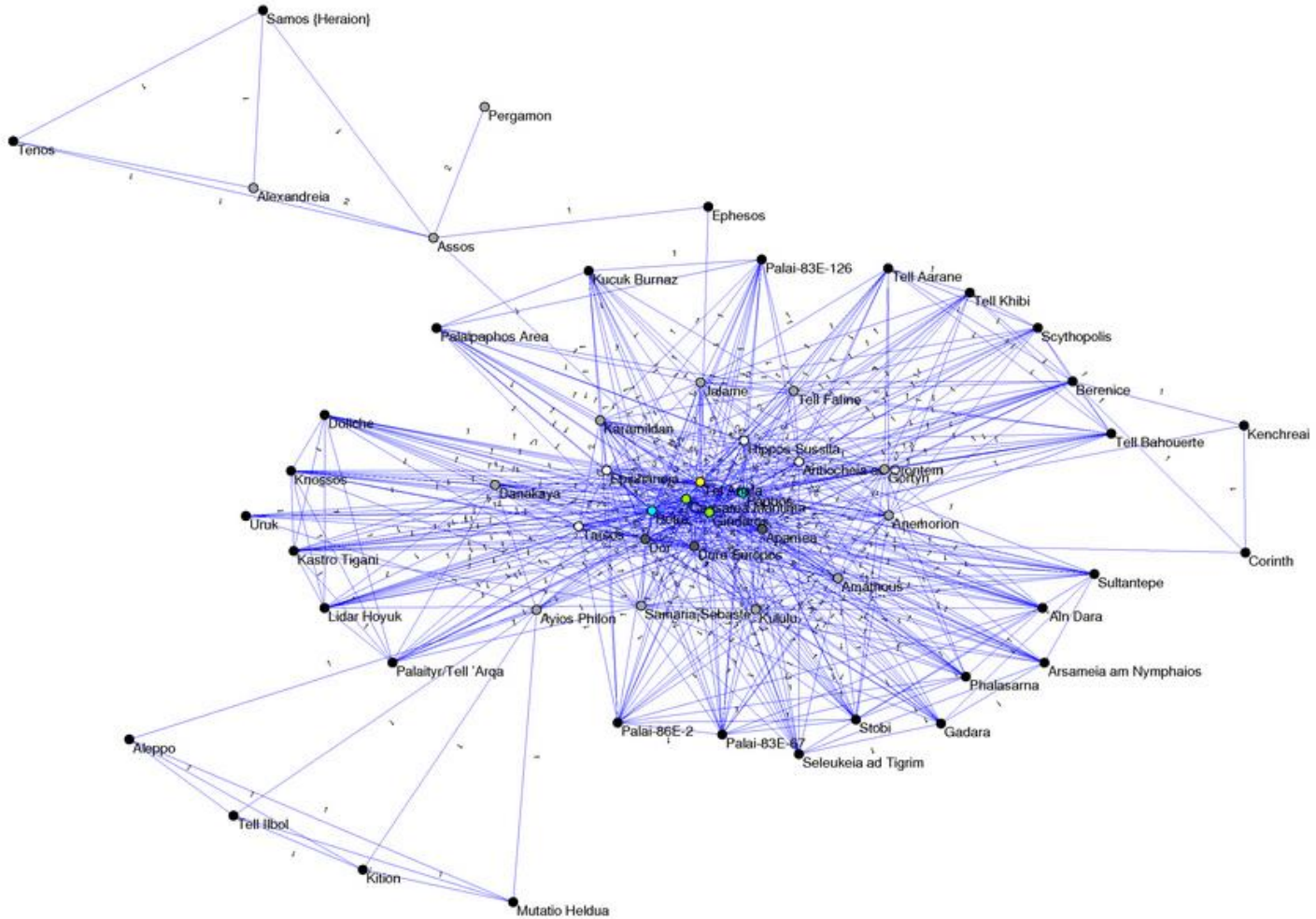
Field Study With Archaeologists

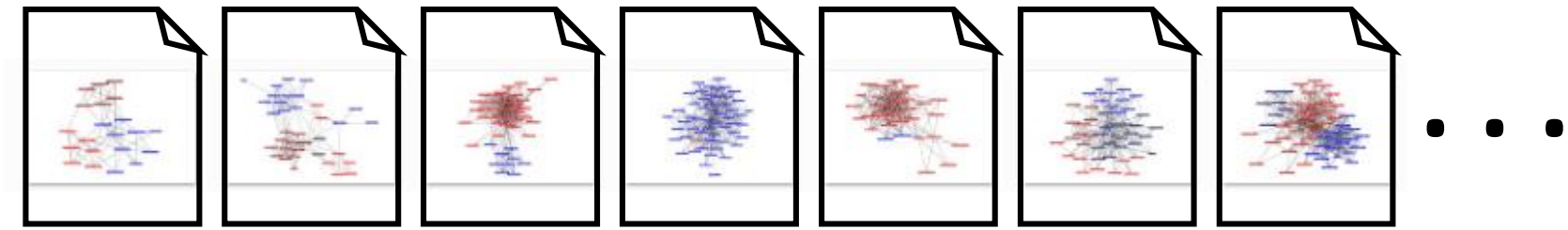
"How were Iron-Age communities integrated into the political and economic structure of the Roman Empire?"

"How were urban social hierarchies within the Roman provinces structured and articulated?"

Brughmans, T., et al., 2011


tbl_iberian_weapons
weapon_id
weapon
town
site_id
reference
inscription
dates
material
period
tbl_location
location
SiteID
tbl_weapons_on_scul...
weapon_sculpture
weapon
town
site_id
reference
dates
tbl_alieni
alieni_id
town
site_id
dating
community
inscription
number_of_inscriptio
ref
note
project_name
unique_ID
tbl_unlimited_visibilit...
unlimited_id
site_from_id
x_coordinate
y_coordinate
tbl_visibility_from
visibility_id
site_from_id
x_coordinate
y_coordinate
tbl_ceramics
SiteID
ceramics_id
ceramics_compor
ceramics_compor
ceramics_compor
ceramics_compor
tbl_site_type
SiteID
Site type
tbl_features
SiteID
Features
tbl_coin_mint
coin_mint_id
site_id
mint_code
mint_description
mint_date
count_coins
tbl_sites
SiteID
unique_ID
site_name
urban_status
elevation
area
x_coordinate
y_coordinate
Municipality
tbl_unlimited_visibilit...
unlimited_id
probability
tbl_visibility_probabil...
visibility_id
probability
tbl_inscriptions
site_ID
traditional_classificati
inscription_ID
dating
tbl_monumentalizati...
monument_id
Site_ID
type
subtype
description
dating
ref
X
Y
code
tbl_dating
SiteID
Dating
tbl_ancient_name
Site_ID
ancient_name
tbl_architecture
SiteID
architecture_id
architecture_compon
architecture_compon
architecture_compon
tbl_unlimited_visibili...
unlimited_id
site_to_id
x_coordinate
y_coordinate
tbl_visibility_to
visibility_id
site_to_id
x_coordinate
y_coordinate
tbl_euergetism
site_ID
EU_ID
Pajek_ID
EUtype
EUmeaning
unique_id
tbl_inscriptions_ID
inscription_id
tbl_inscriptions_dates
inscription_id
dating
tbl_monumentalizati...
monument_id
project_period


Samos (Heraion)
Pergamon
Tenos
Alexandreia
Assos
Ephesos
Kucuk Burnaz
Palai-83E-126
Tell Aarane
Tell Khibi
Scythopolis
Palaipaphos Area
Jalame
Tell Afine
Berenice
Doliche
Karamidan
Tell Bahoueite
Kenchreai
Knossos
Danakaya
Uruk
Anemorion
Kastro Tigani
Corinth
Sultantepe
Amathous
Lidar Hoyuk
Ayios Philon
Ain Dara
Palaityr/Tell 'Arqa
Arsameia am Nymphaios
Phalasarna
Palai-86E-2
Palai-83E-67
Stobi
Gadara
Seleukeia ad Tigrim
Aleppo
Tell Ilbol
Kition
Mutatio Heldua




Sites
Inscriptions
Sites
Alamillo Alcalá del Río Alcalá Morisco
Alcaudete Alcolea del Río
Almadén de la Plata Almodóvar del Río
Aparicio el Grande Arcos de la Frontera
Atalaya de la Morenilla Azmalcozar
Bonanza Caganche
Cañada Real de Morón (IX)
Cañada Real de Morón (VII)
Cañada Real de Morón (VIII)
Cañada Real de Morón (XII) Cantillana
Cepaperra Carmona
Aliens
Baetica
Sites
Inscriptions
Inscriptions
Aliens
Baetica
Aliens
Uncertain
Aliens
Hispania
Aliens
Beyond
Sites
Sites
Sites
Sites
Sites
Expand Site to Aliens
Sites

GraphTrail

Scalability

1. Number of nodes, edges, types

	Nodes	Types	Edges	Types
CHI	10K	3	20K	3+
Archaeology	13K	24	20K	35

2. Number of charts

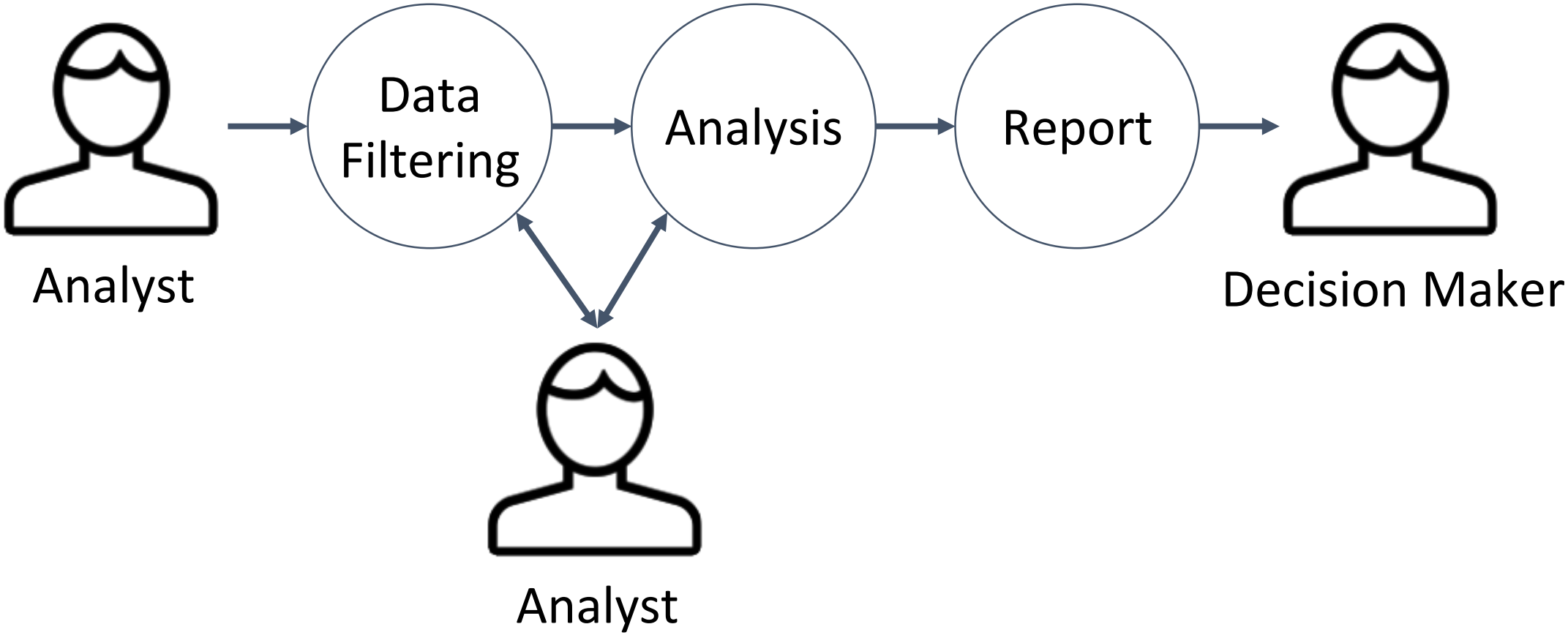
   20 – 30 per session

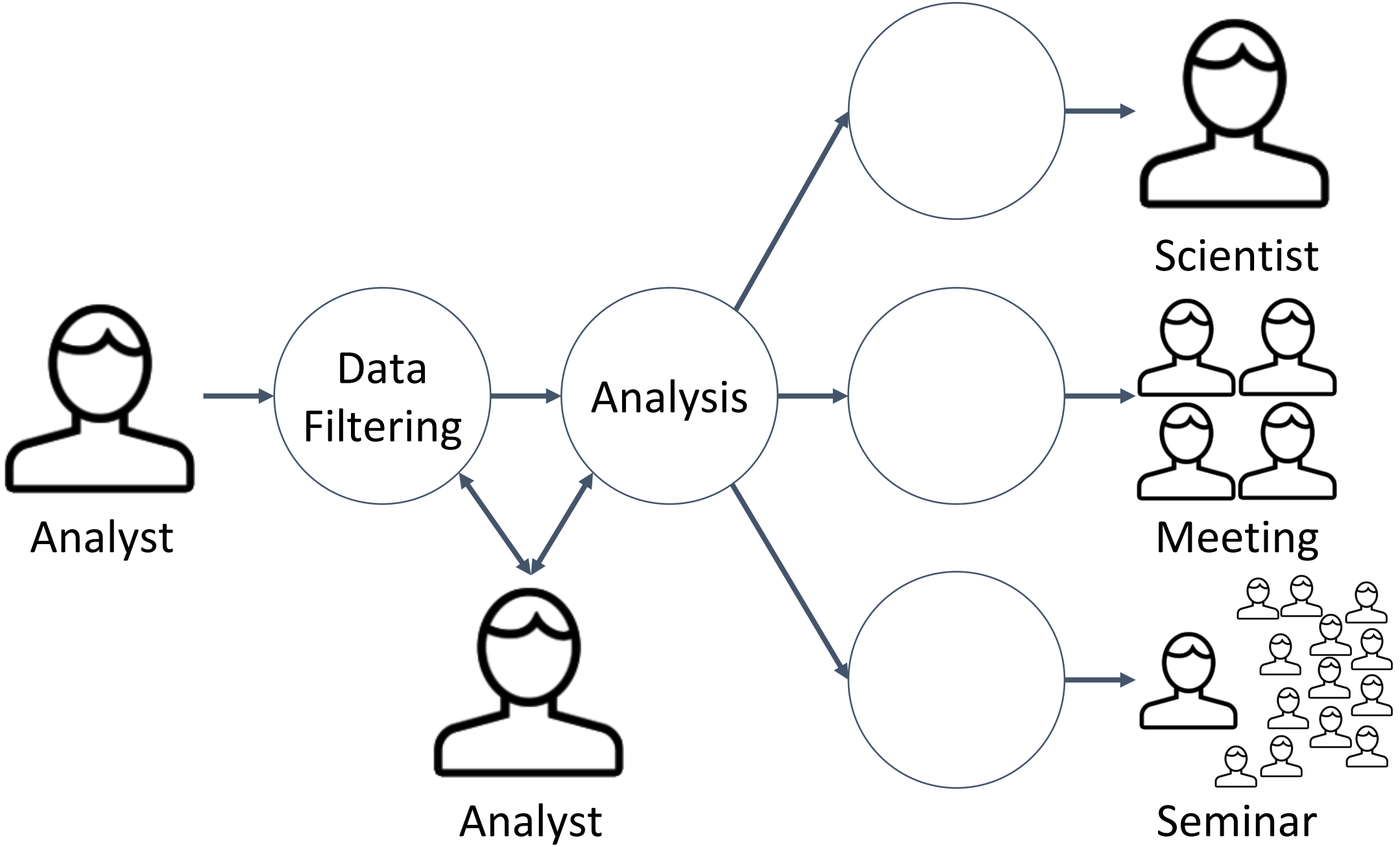
GraphTrail

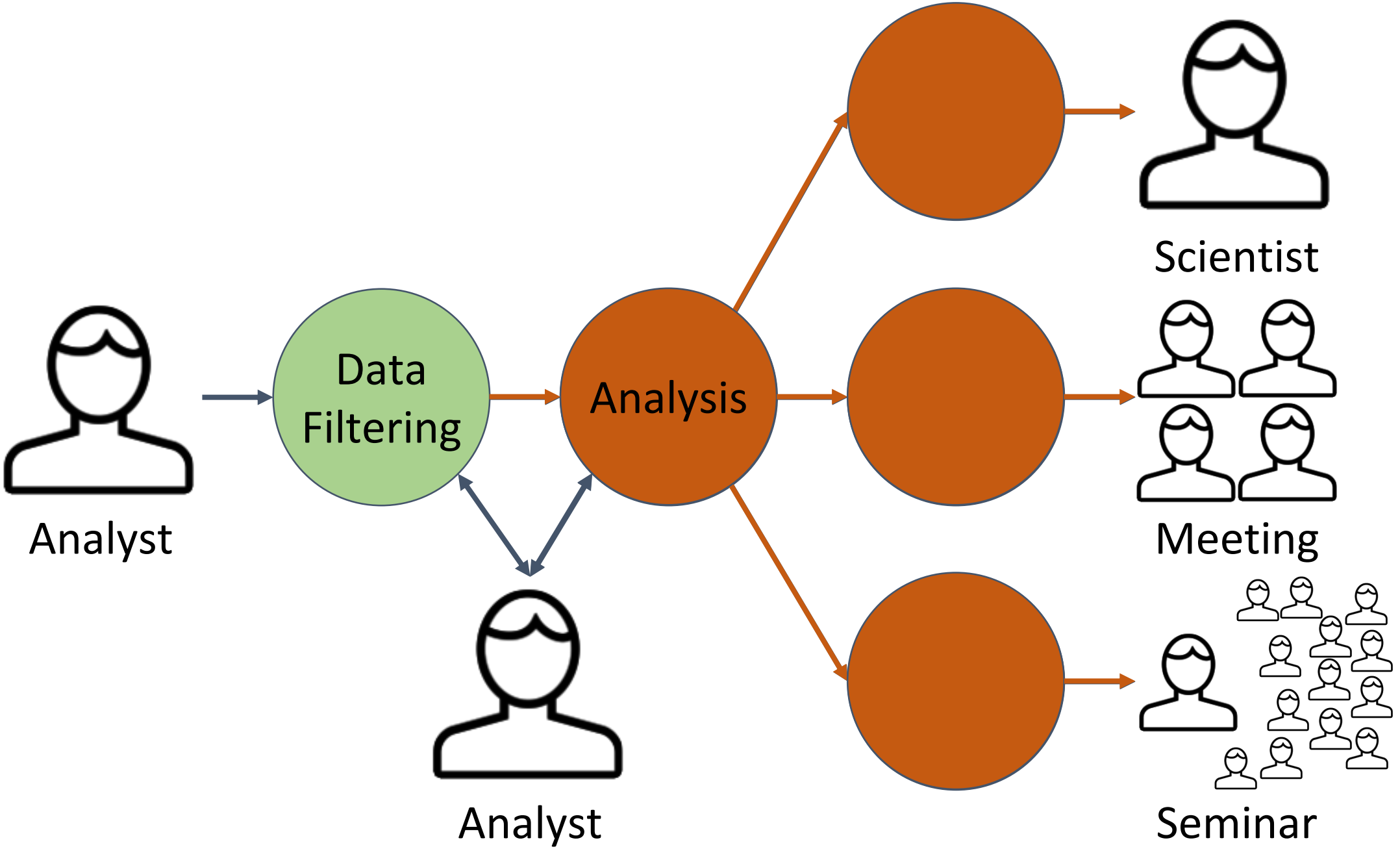
- A system for exploring **large multivariate**, **heterogeneous networks** using **aggregation** by node and edge attributes,
- A method for capturing **data provenance** and integrating it directly into the workspace, and
- A longitudinal **field study** and a qualitative **lab study** that demonstrate the utility of these approaches.

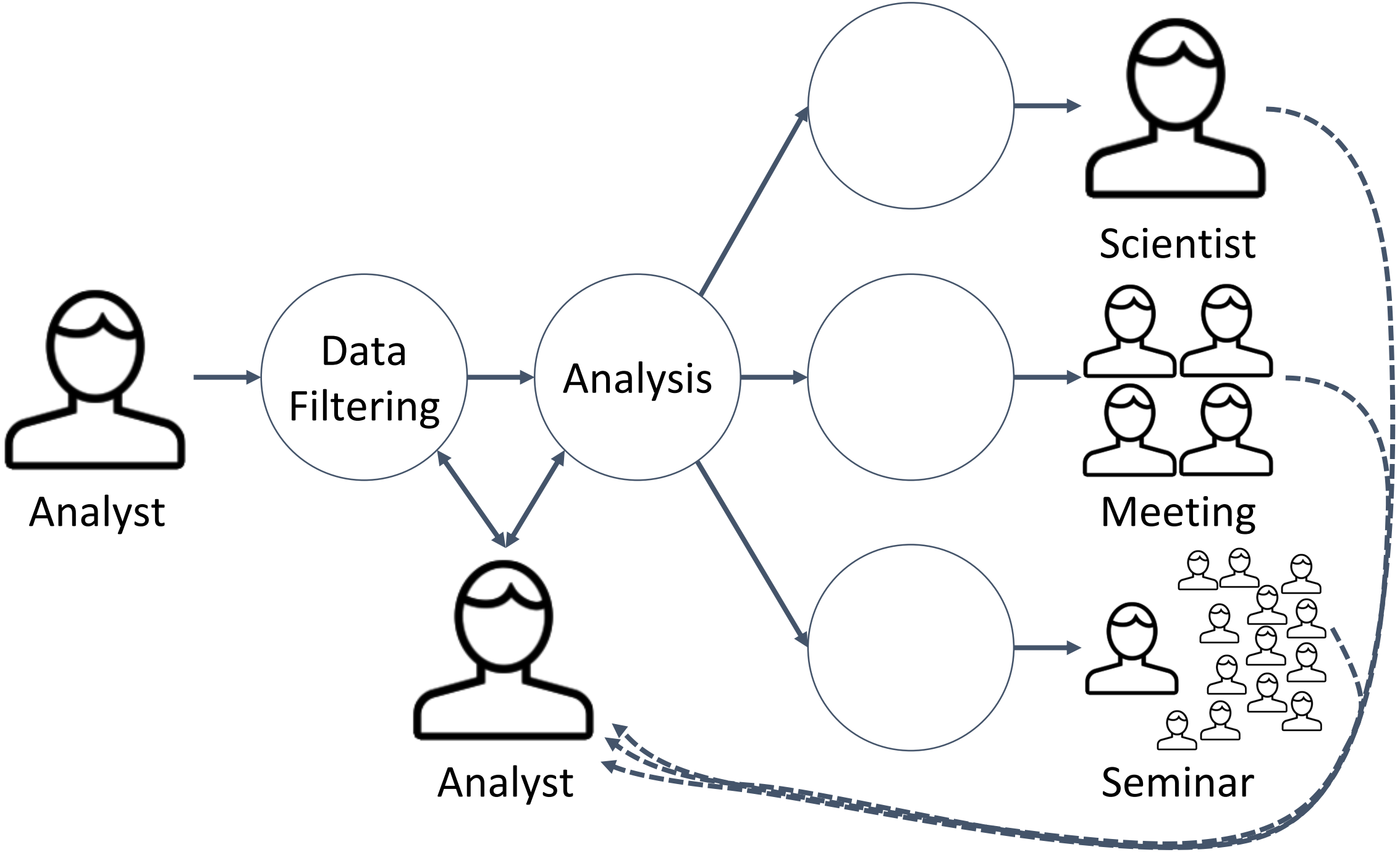
Dunne C, Riche NH, Lee B, Metoyer RA and Robertson GG (2012), *"GraphTrail: Analyzing large multivariate, heterogeneous networks while supporting exploration history"*, In CHI `12.

Riche N, Lee B and Dunne C (2011), *"Interactive visualization for exploring multi-modal, multi-relational, and multivariate graph data"*. US Patent No. (8743122).

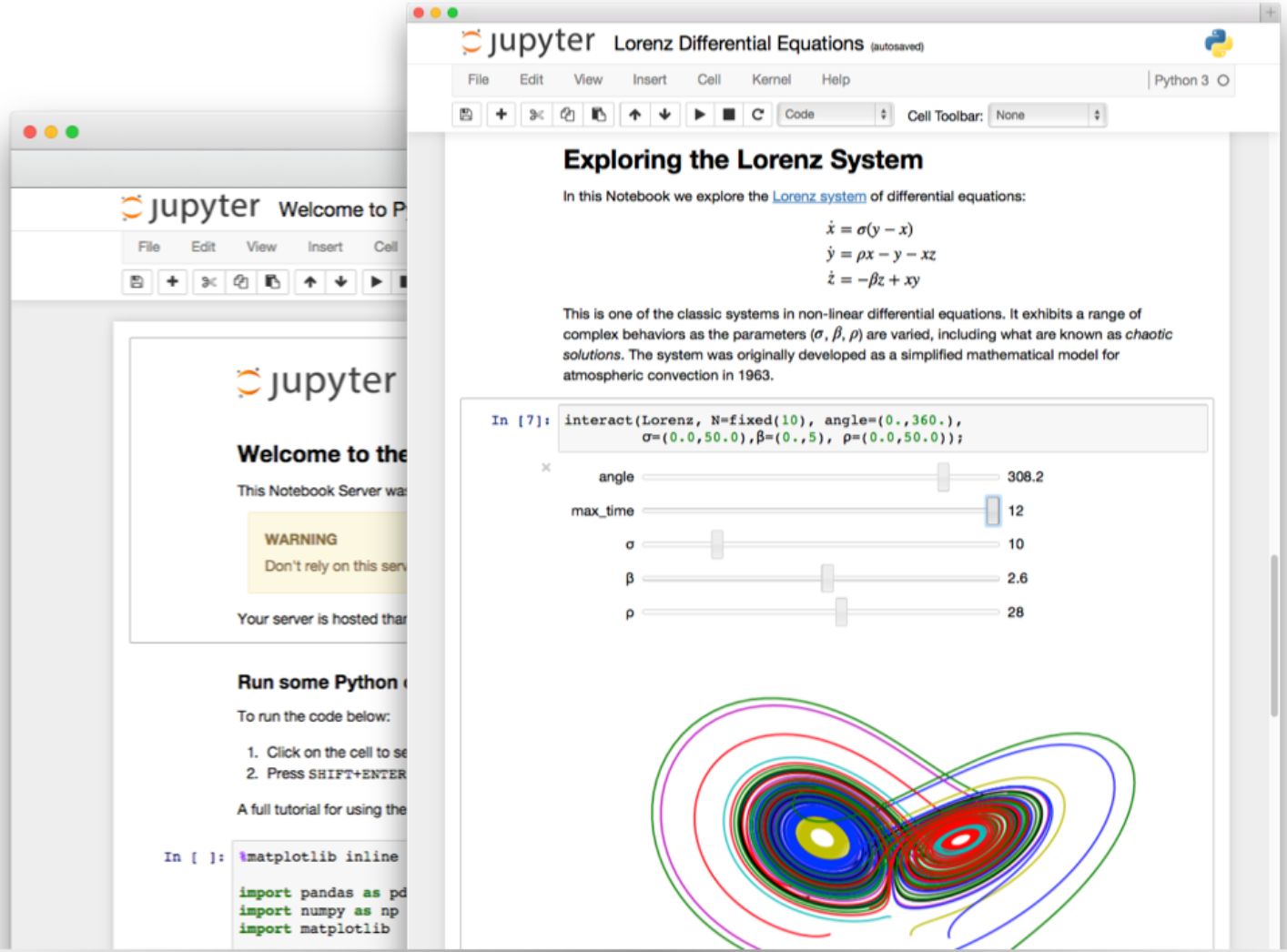

Data
Filtering
Analysis
Report
Analyst
Decision Maker
Analyst


Analyst
Data Filtering
Analysis
Analyst
Scientist
Meeting
Seminar


Analyst
Data Filtering
Analysis
Analyst
Scientist
Meeting
Seminar


Analyst
Data Filtering
Analysis
Scientist
Meeting
Seminar
Analyst

CENTER FOR
OPEN SCIENCE



Films

Films	
title	string
episode_id	integer
opening_crawl	string
director	string
producer	string
release_date	date
species	array
starships	array
vehicles	array
characters	array
planets	array
url	string
created	string
edited	string

Starships	
name	string
model	string
starship_class	string
manufacturer	string
cost_in_credits	string
length	string
crew	string
passengers	string
max_atmosphere_speed	string
hyperdrive_rating	string
MGLT	string
cargo_capacity	string
consumables	string
films	array
pilots	array
url	string
created	string
edited	string

Starships

Species	
name	string
classfication	string
average_height	Type
average_lifespan	string
eye_colors	string
hair_colors	string
skin_colors	string
language	string
homeworld	string
people	array
films	array
url	string
created	string
edited	string

Species

People

People	
name	string
birth_year	string
eye_color	string
gender	string
hair_color	string
height	string
mass	string
skin_color	string
homeworld	string
films	array
species	array
starships	array
vehicles	array
url	string
created	string
edited	string

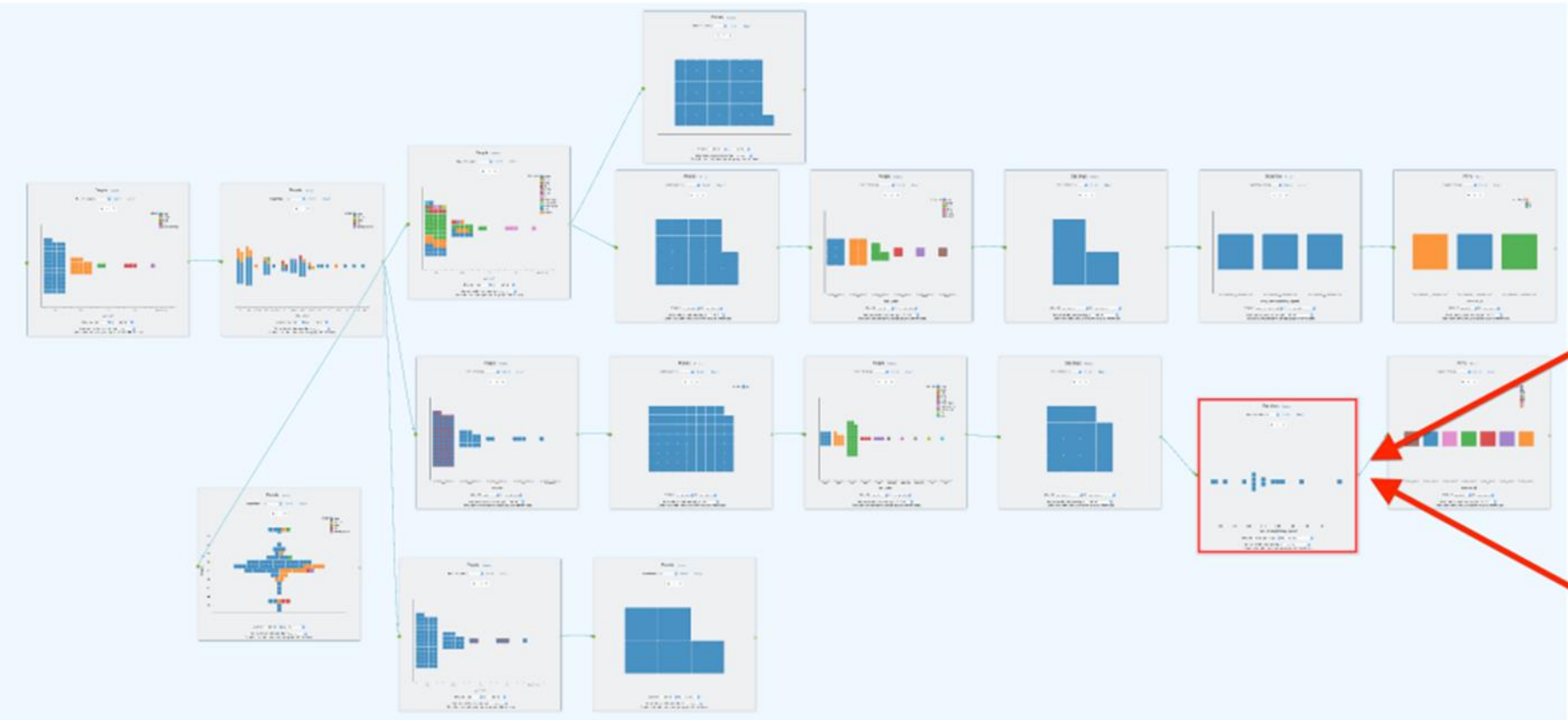
Vehicles

Vehicles	
name	string
model	string
vehicle_class	string
manufacturer	string
length	string
cost_in_credits	string
crew	string
passengers	string
max_atmosphering_speed	string
cargo_capacity	string
consumables	string
films	array
pilots	array
url	string
created	string
edited	string

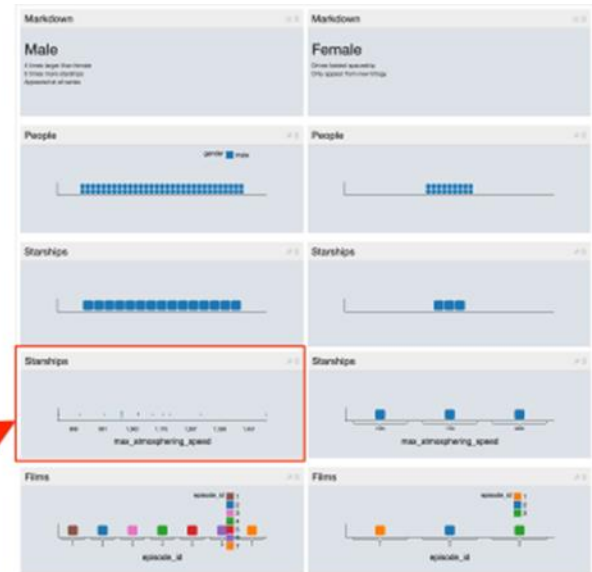
Planets	
name	string
diameter	string
rotation_speed	string
orbital_period	string
gravity	string
population	string
climate	string
terrain	string
surface_water	string
residents	array
films	array
url	string
created	string
edited	string

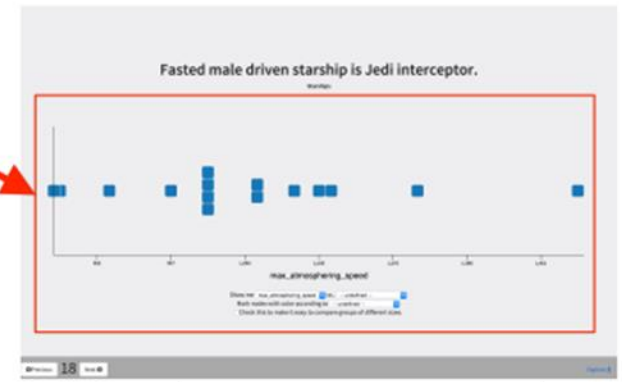
Planets

Trail



Dashboard

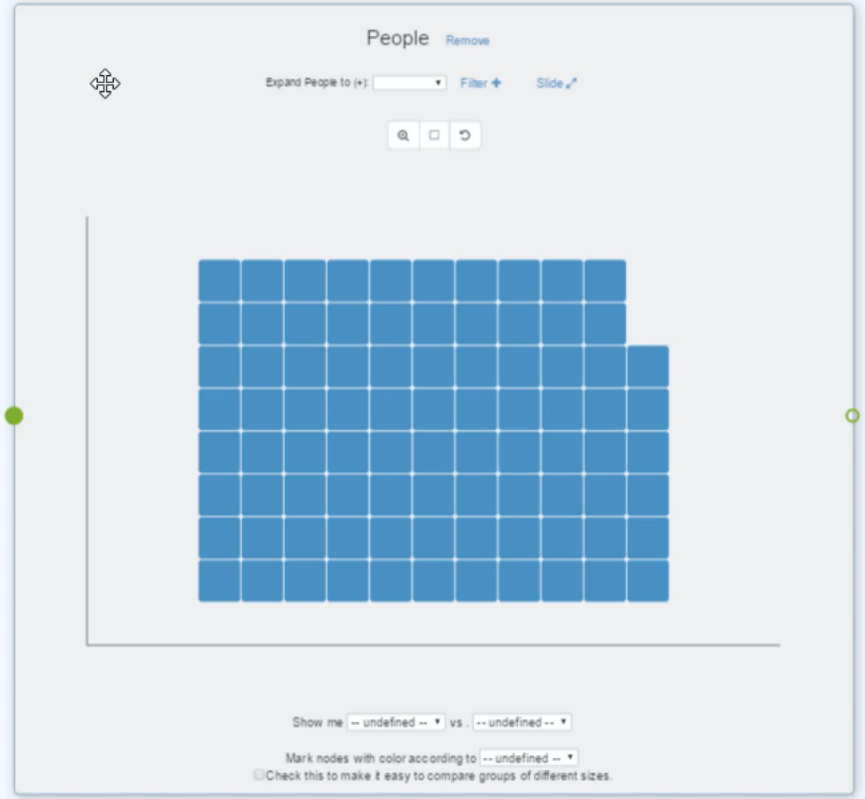




Slideshow

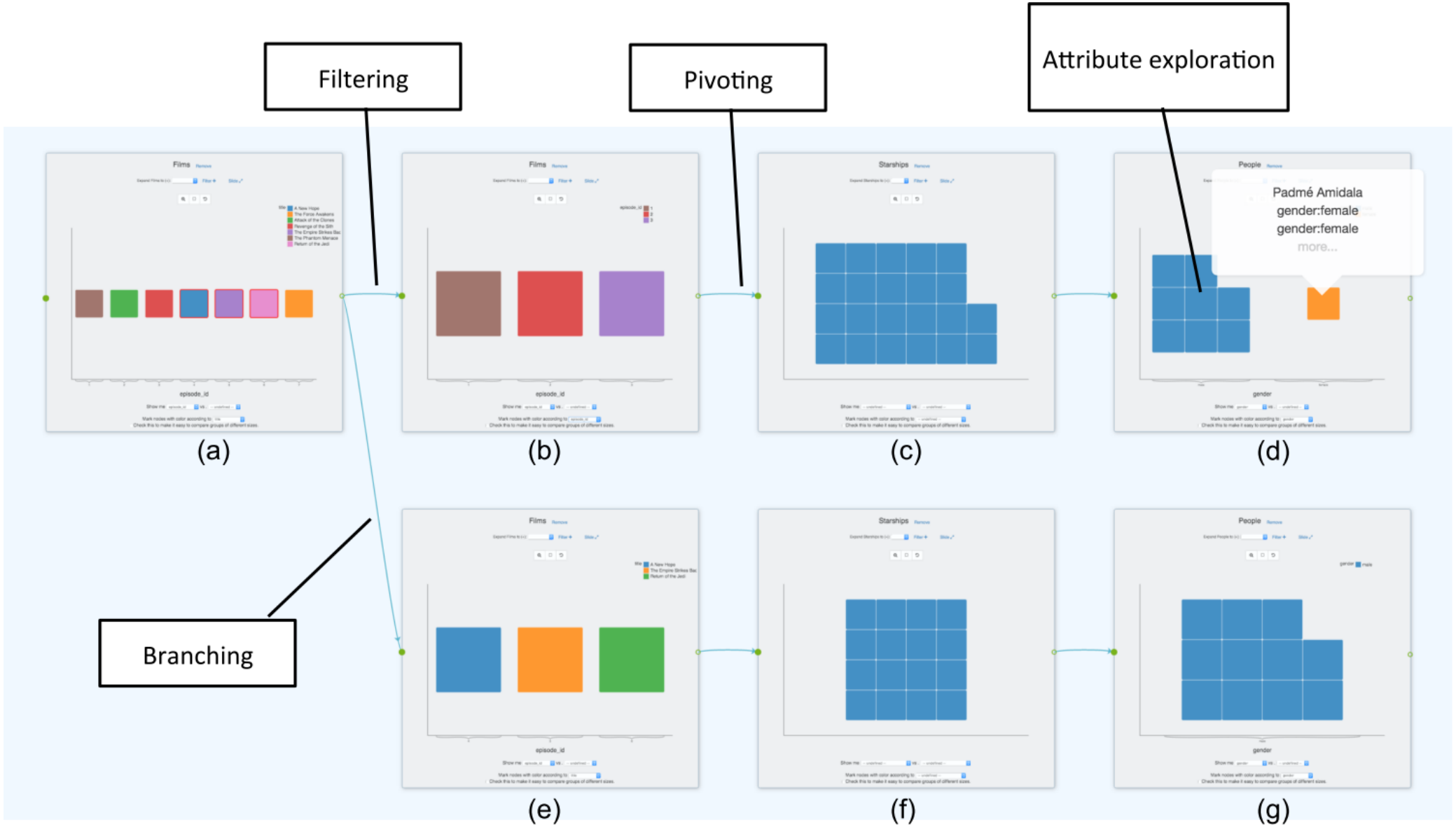
Entity

Public



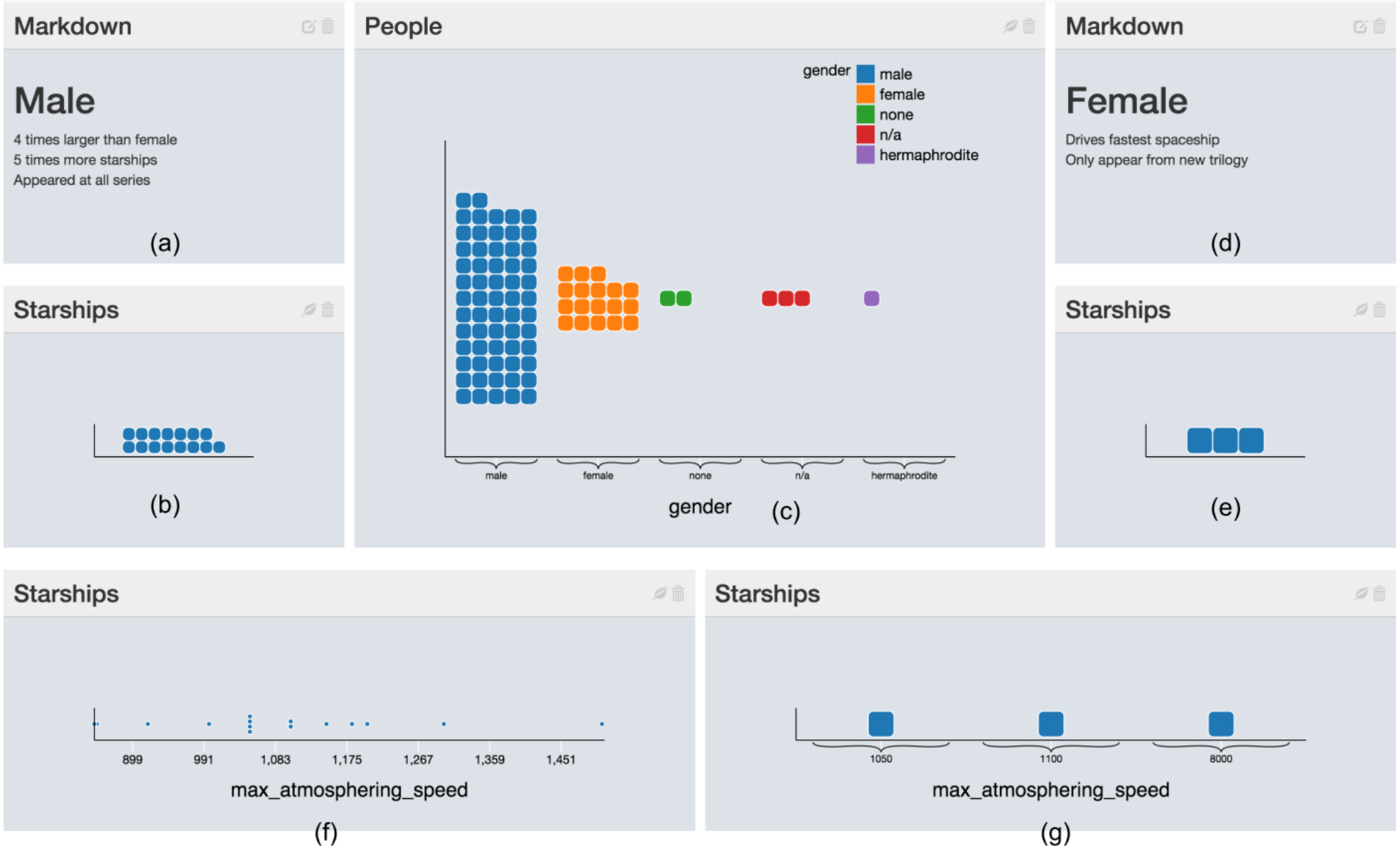
StoryFacets

Trail facet – Star Wars trilogy ships & pilots



StoryFacets

Dashboard facet – Star Wars male vs. female



StoryFacets

Infographic facet – Star Wars Jabba the Hutt

Markdown

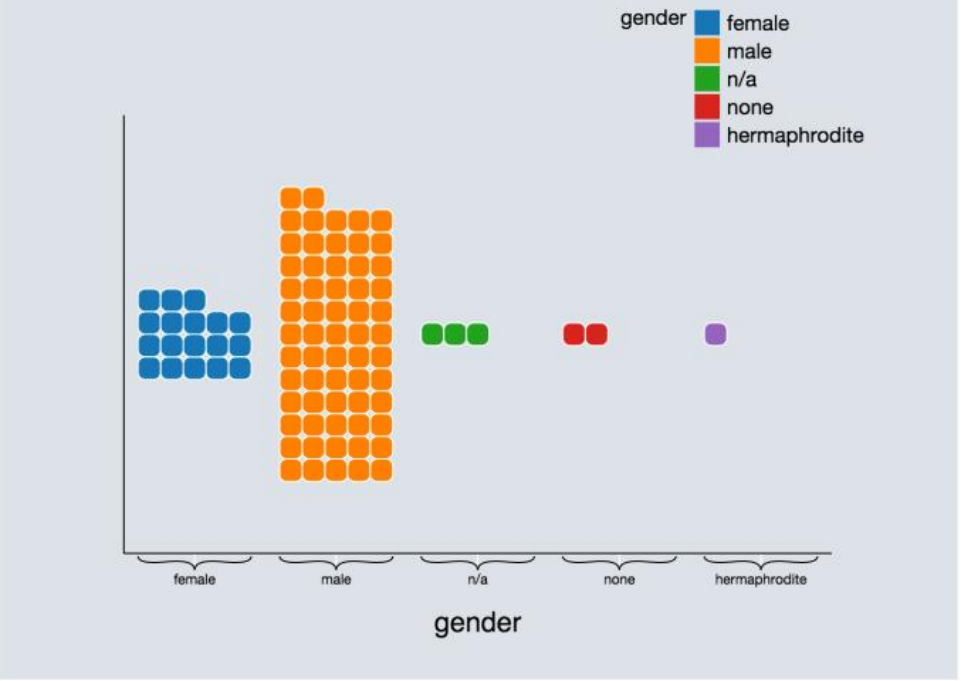
(a)

Markdown

The only hermaphrodite in the Starwars is Jabba Desilijic Tiure

(b)

People



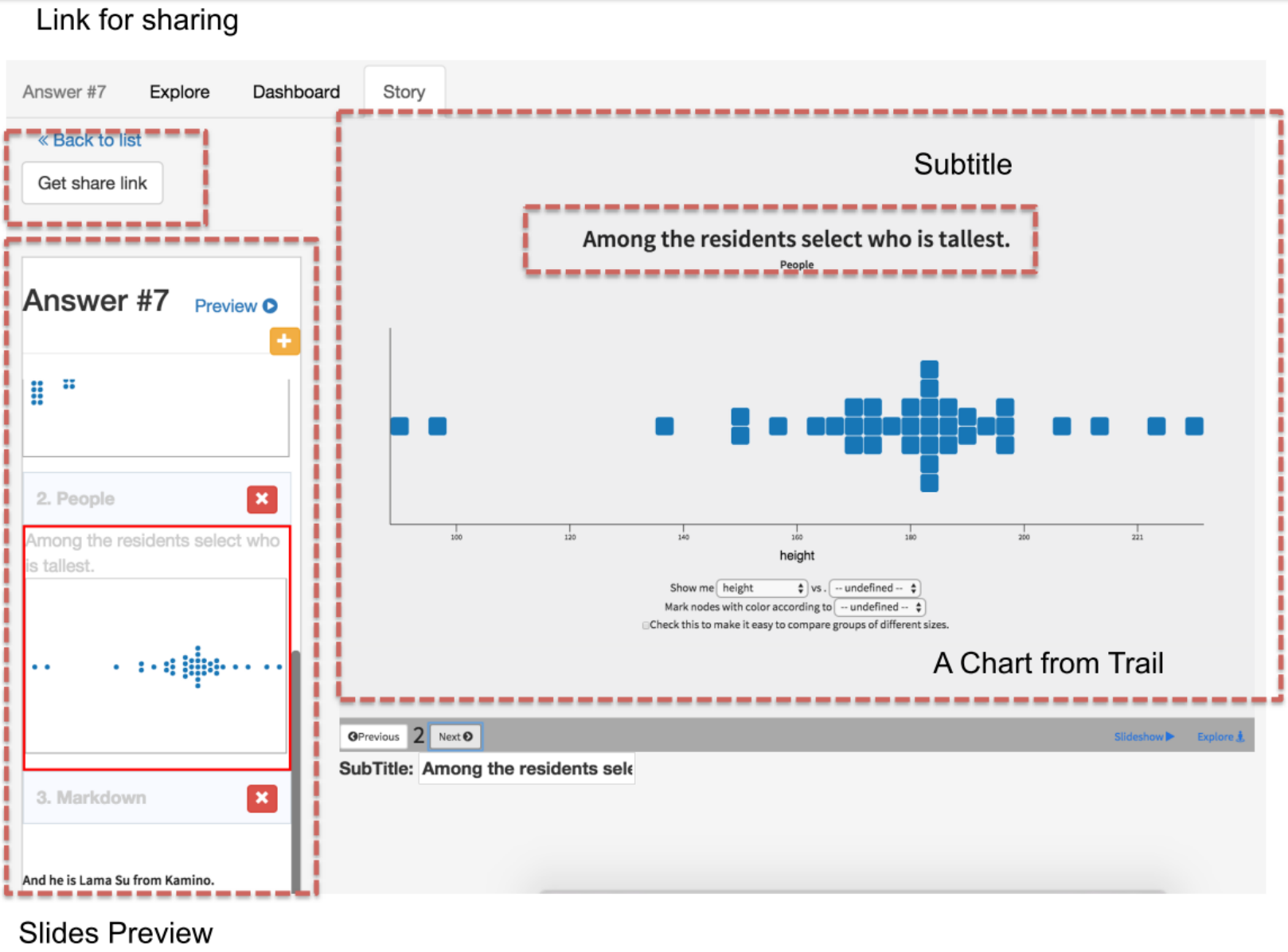
(c)

Markdown

(d)

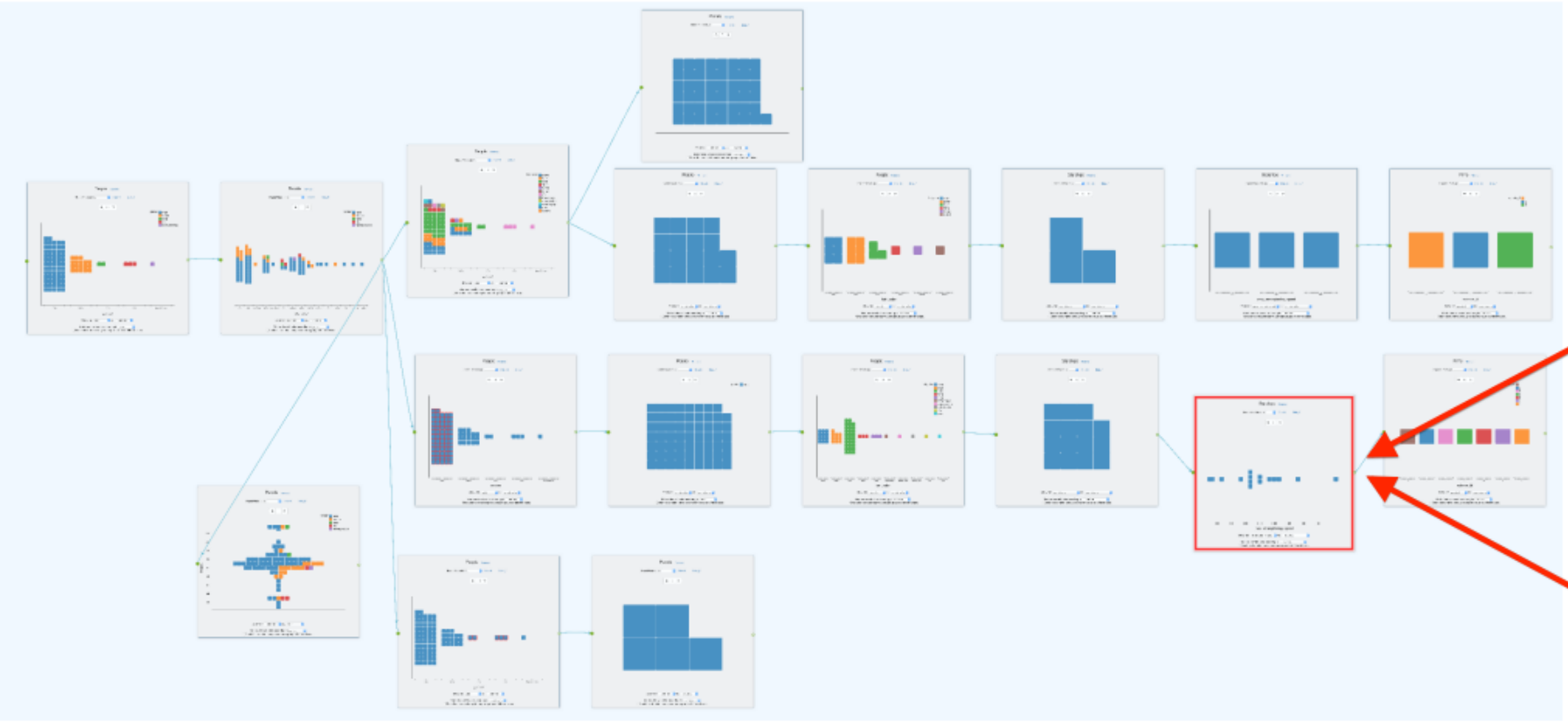
StoryFacets

Story (slideshow) facet – Star Wars character height



StoryFacets

Linked back to trail facet



(a) Trail Facet

(b) Dashboards or Info Graphics Facet

(c) Story Facet

Evaluation

19 casual participants in a public setting

- Novice users select different communication formats based on target audience and nature of the content

Three visualization professional expert reviewers

- Different & specific uses of the three views

Evaluate how view affects presentation style and structure

Future: Evaluate how view, context, and expertise affect knowledge transfer

StoryFacets

- Exploratory data analysis is much more than the initial exploration session
- Unified platform to support exploration, collaboration, discussion, and presentation
- Current & ongoing studies reveal future directions

Dunne C, Skelton C, Diamond S, Martino M and Meirelles I (2016), *"Quantitative, qualitative, and historical urban data visualization tools for professionals and stakeholders"*, In HCI International `16.

Park DG, Dunne C, Ragan E, Elmqvist N (2016) *"StoryFacets: Generating Multiple Representations of Exploratory Data Analysis for Communication",* Under submission.

GraphTrail & StoryFacets

Blocks to build on

- Additional visualizations
- Context aware comments
- Exploration hints for new paths
- Streaming/temporal data, intelligent updates, and resurgent relevancy
- Linked chart parameterization for comparisons with auto chain layout and compression
- Advanced modeling and analytics (IBM Catalyst, IBM Watson)
- User management & security