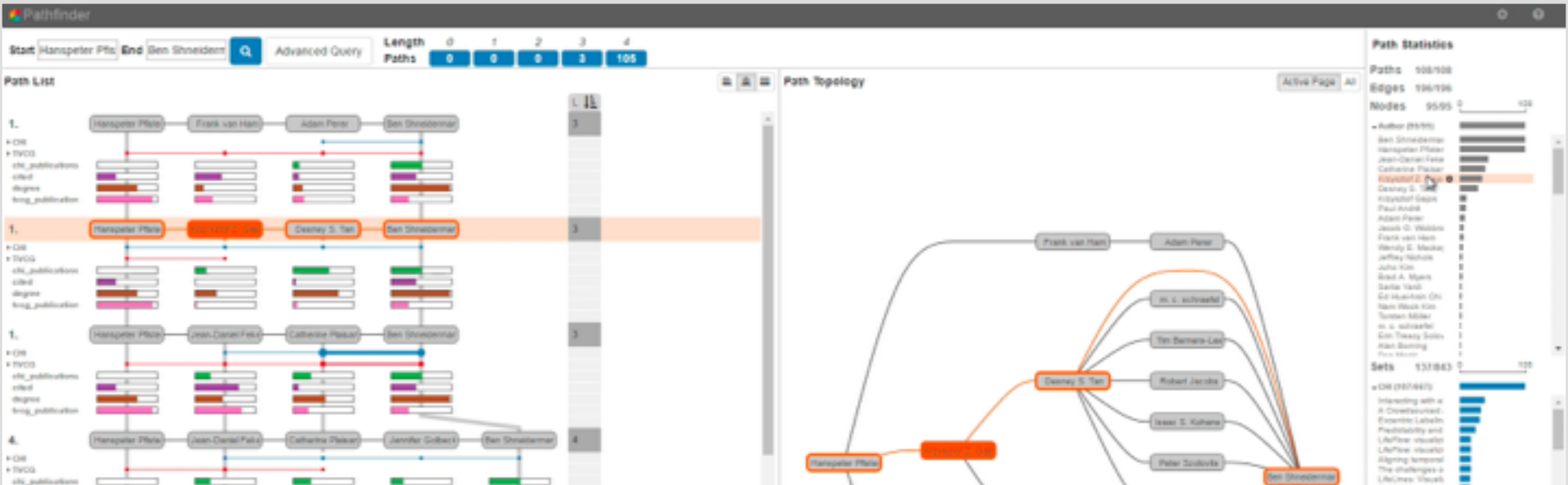
Pathfinder: Visual Analysis of Paths in Graphs

Christian Partl, Samuel Gratzl, Marc Streit, Anne Mai Wassermann,
Hanspeter Pfister, Dieter Schmalstieg, **Alexander Lex**

CALEYDO

visualization design lab

THE UNIVERSITY OF UTAH

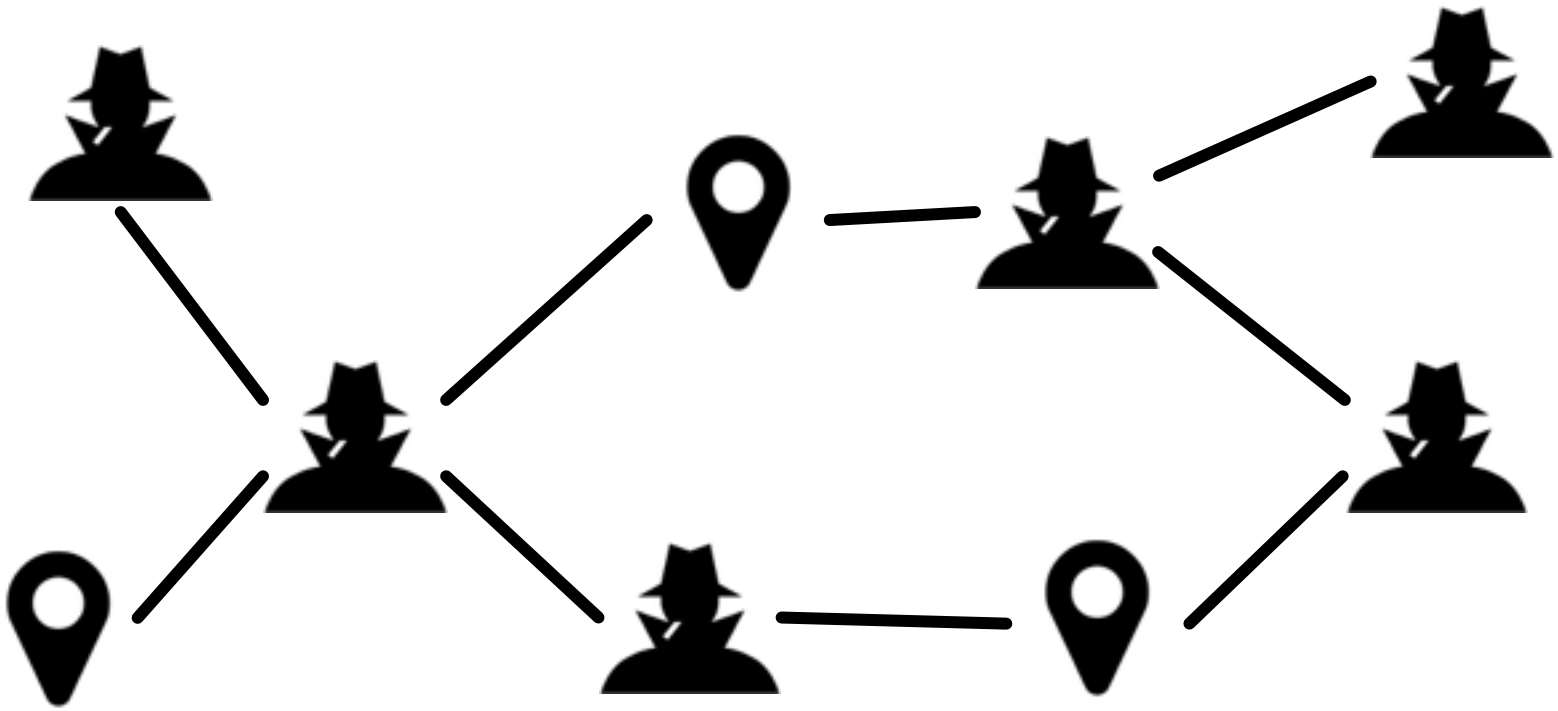


Tool / Paper

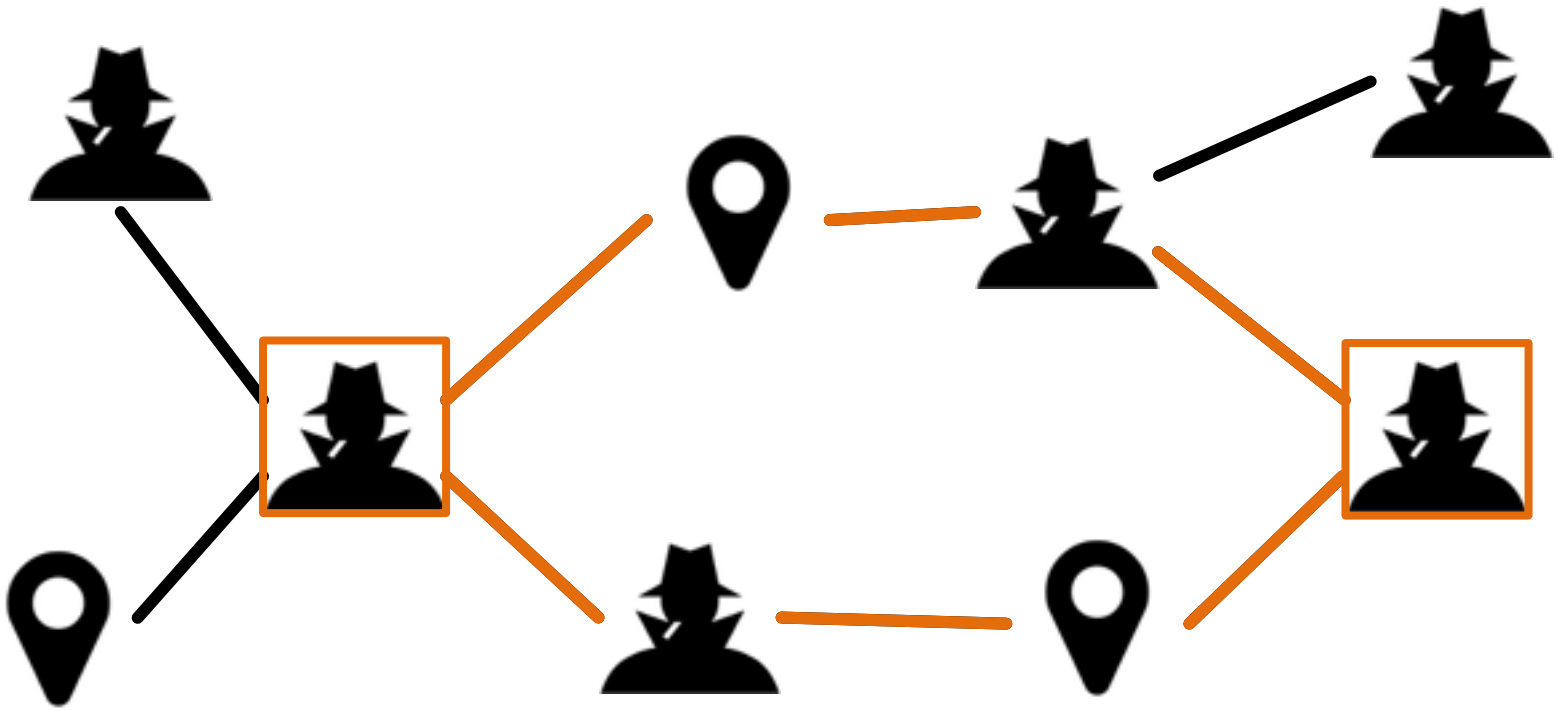
Christian Partl, Samuel Gratzl, Marc Streit, Anne Mai Wassermann, Hanspeter Pfister, Dieter Schmalstieg, Alexander Lex, **Pathfinder: Visual Analysis of Paths in Graphs,** Computer Graphics Forum (EuroVis '16) (to appear), 2016.

Tool: http://demo.caleydo.org/pathfinder/

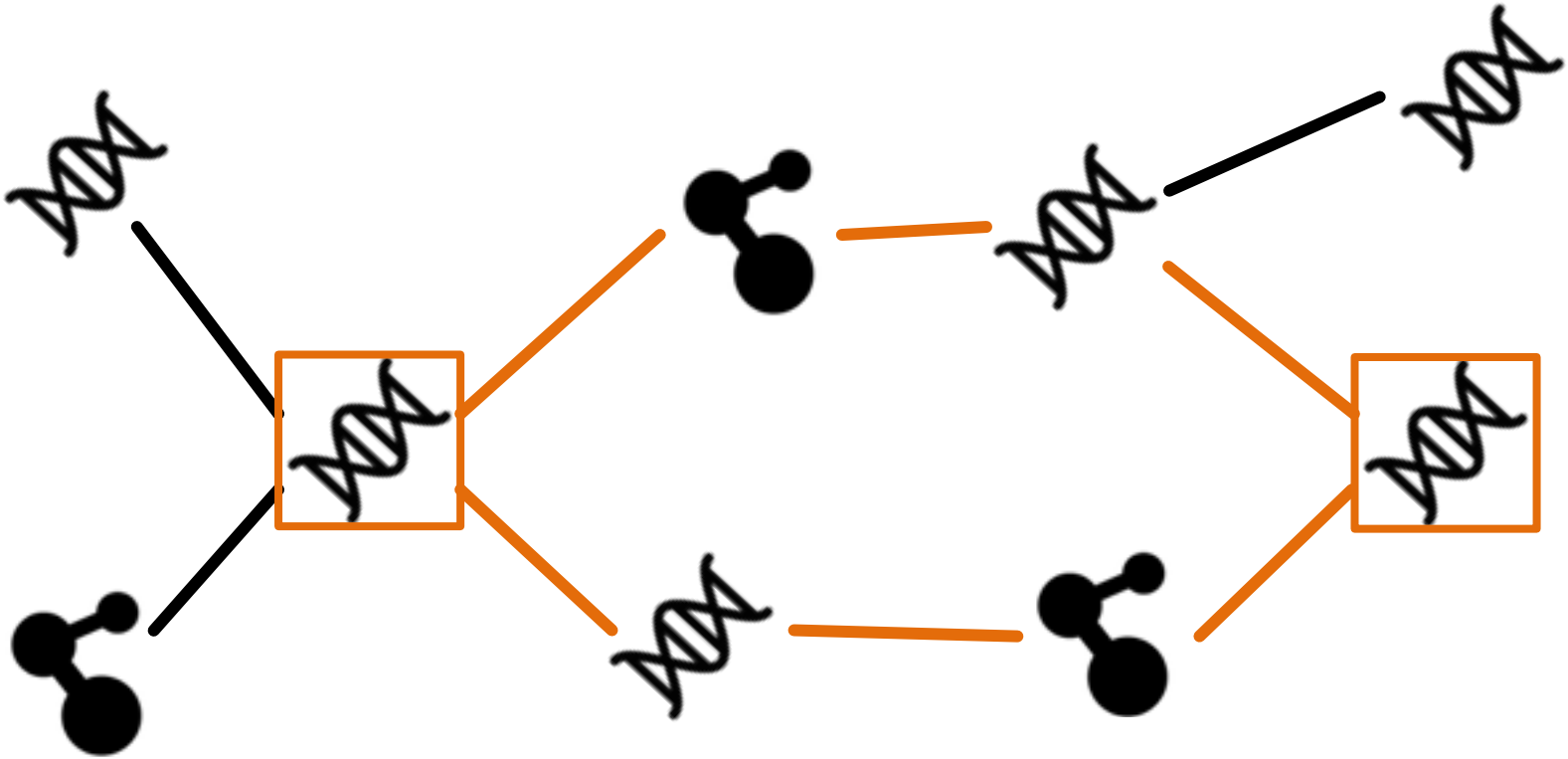
Info: http://www.caleydo.org/publications/2016_eurovis_pathfinder/



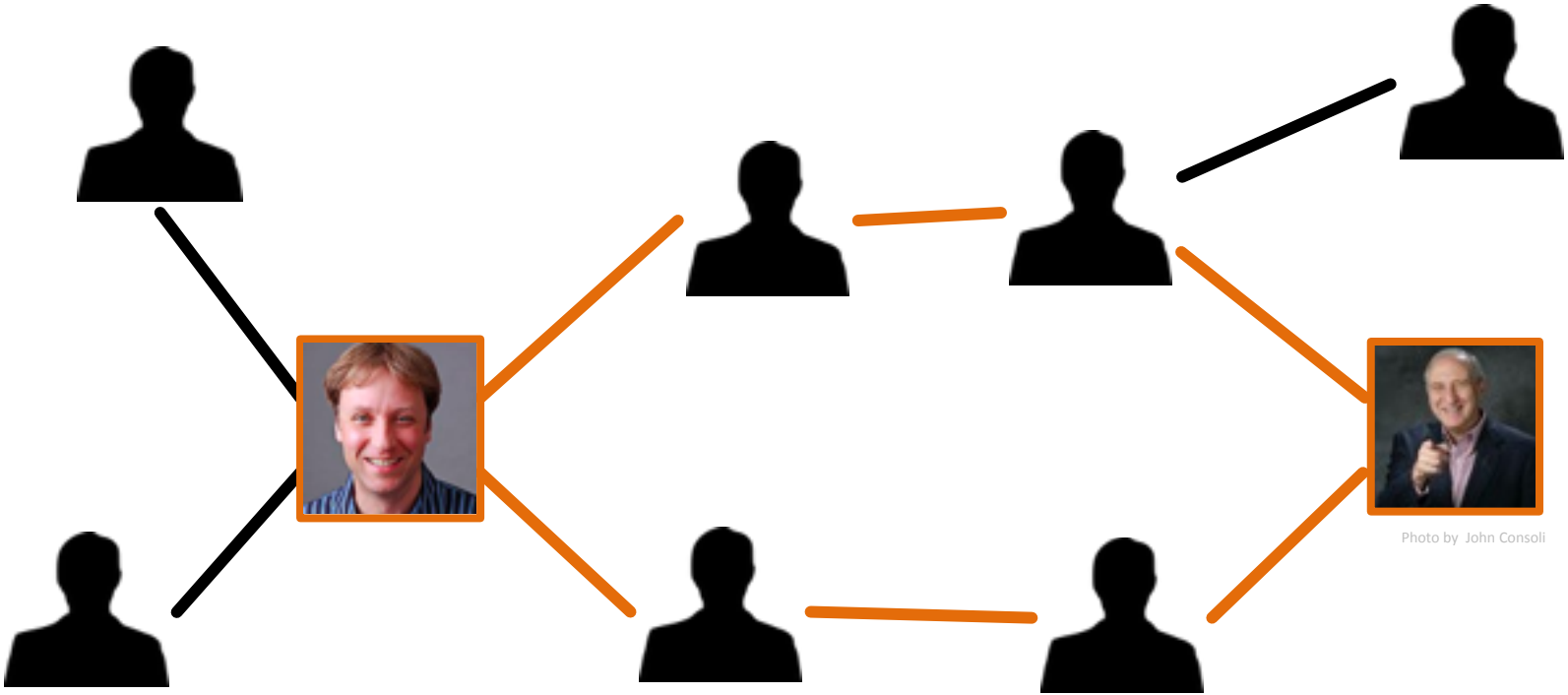
Intelligence Data: How are two suspects connected?



Intelligence Data: How are two suspects connected?

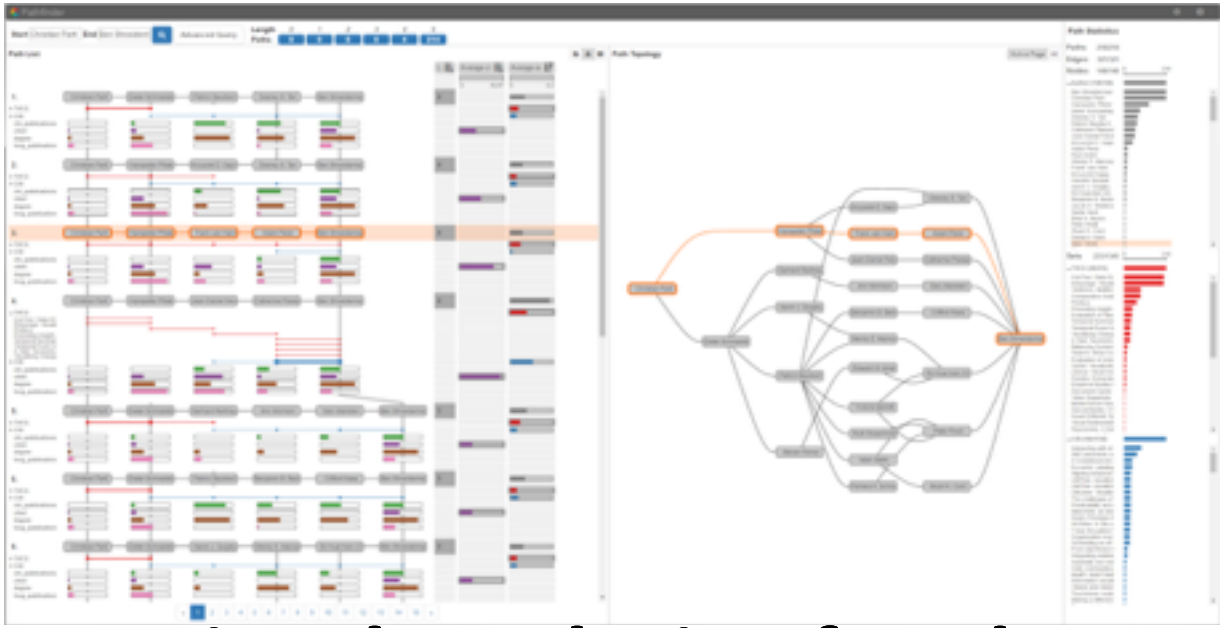


Biological Network: How do two genes interact?



Coauthor Network: How is HP Pfister connected to Ben Shneiderman?

Pathfinder



Visual Analysis of Paths

in Large Multivariate Graphs

Challenge 1: Graph Size

How can we deal with graphs too large to sensibly render at once?

Approach: **(Path) Queries**

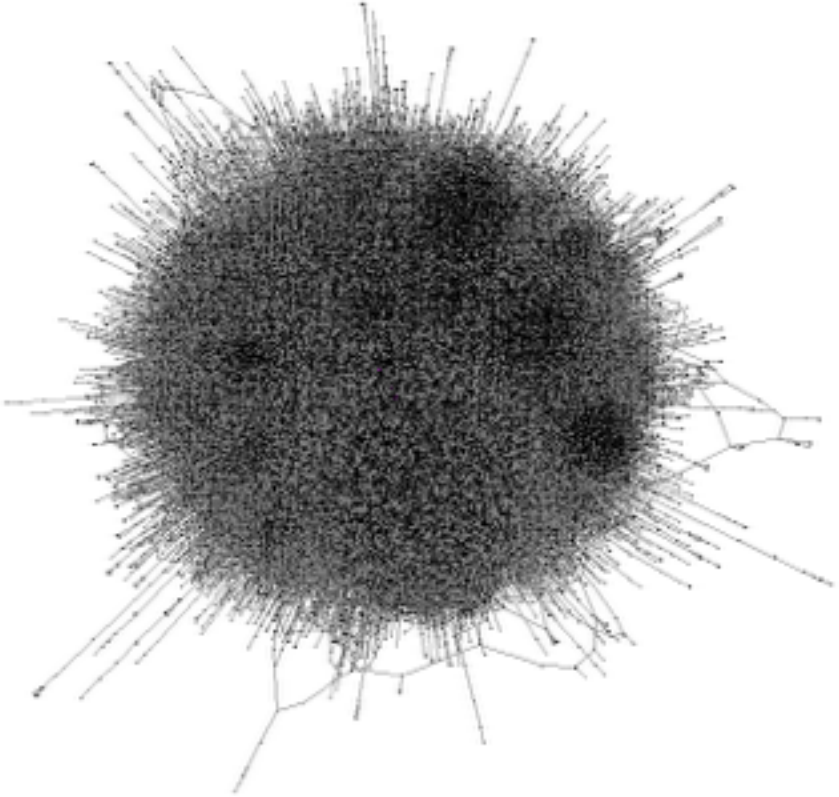
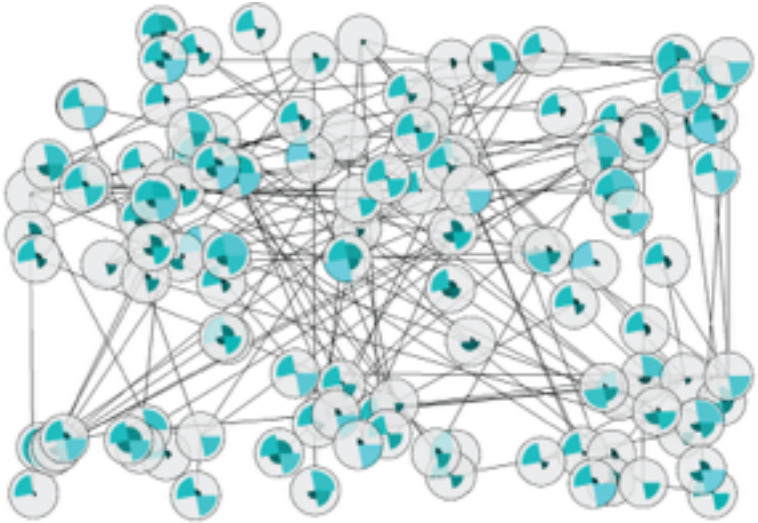


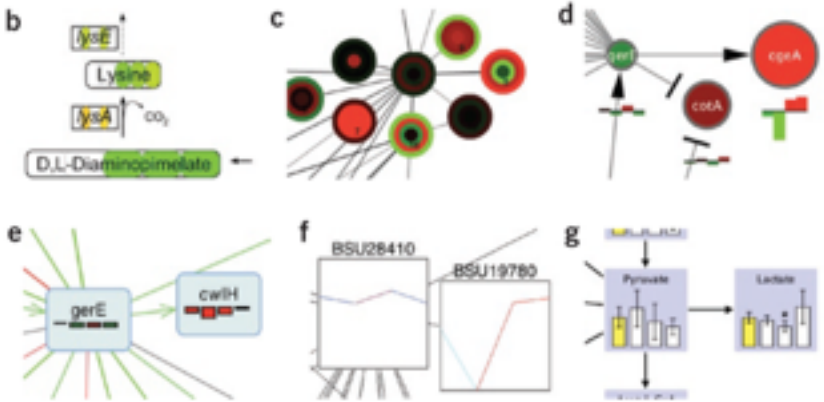
Image Source:
http://www.thenetworkthinkers.com/2013/03/big-data.html

Challenge 2: Multivariate Attributes

How can we deal with graphs that contain **rich attribute data**?



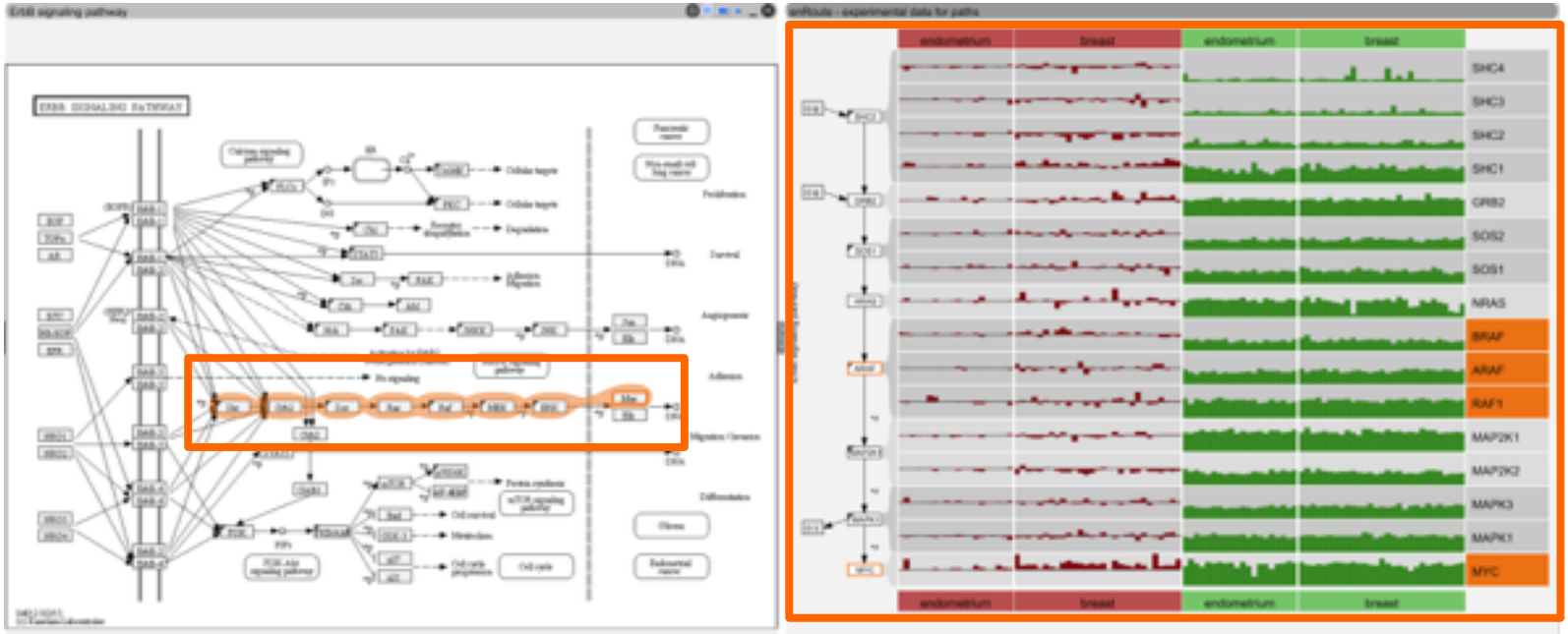
[McDonnel2009]



[Gehlenborg2010]

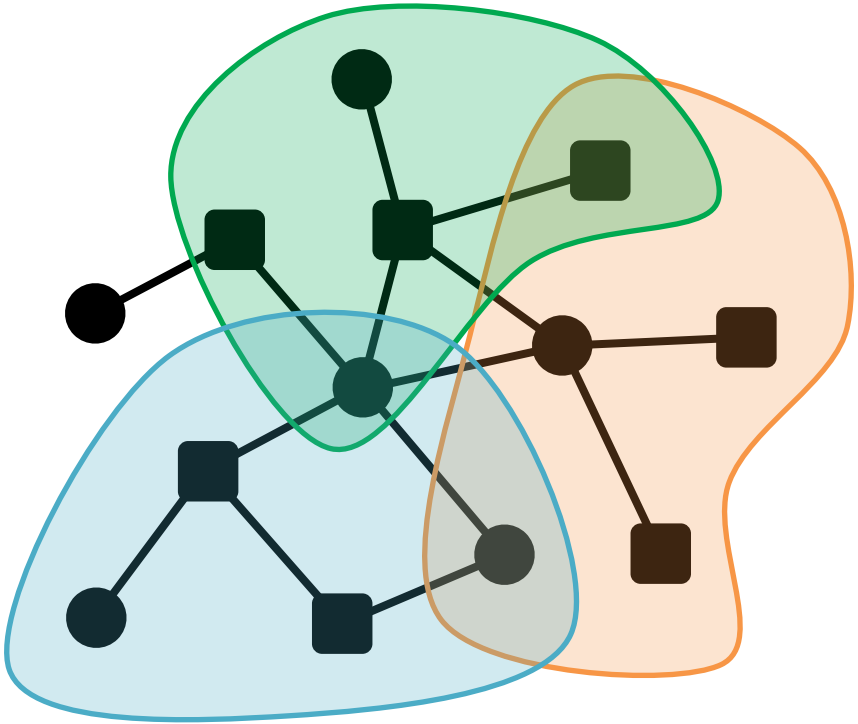
Approach: Attribute Visualization for Paths

EnRoute http://enroute.caleydo.org



[Partl2012]

Network Data



Nodes

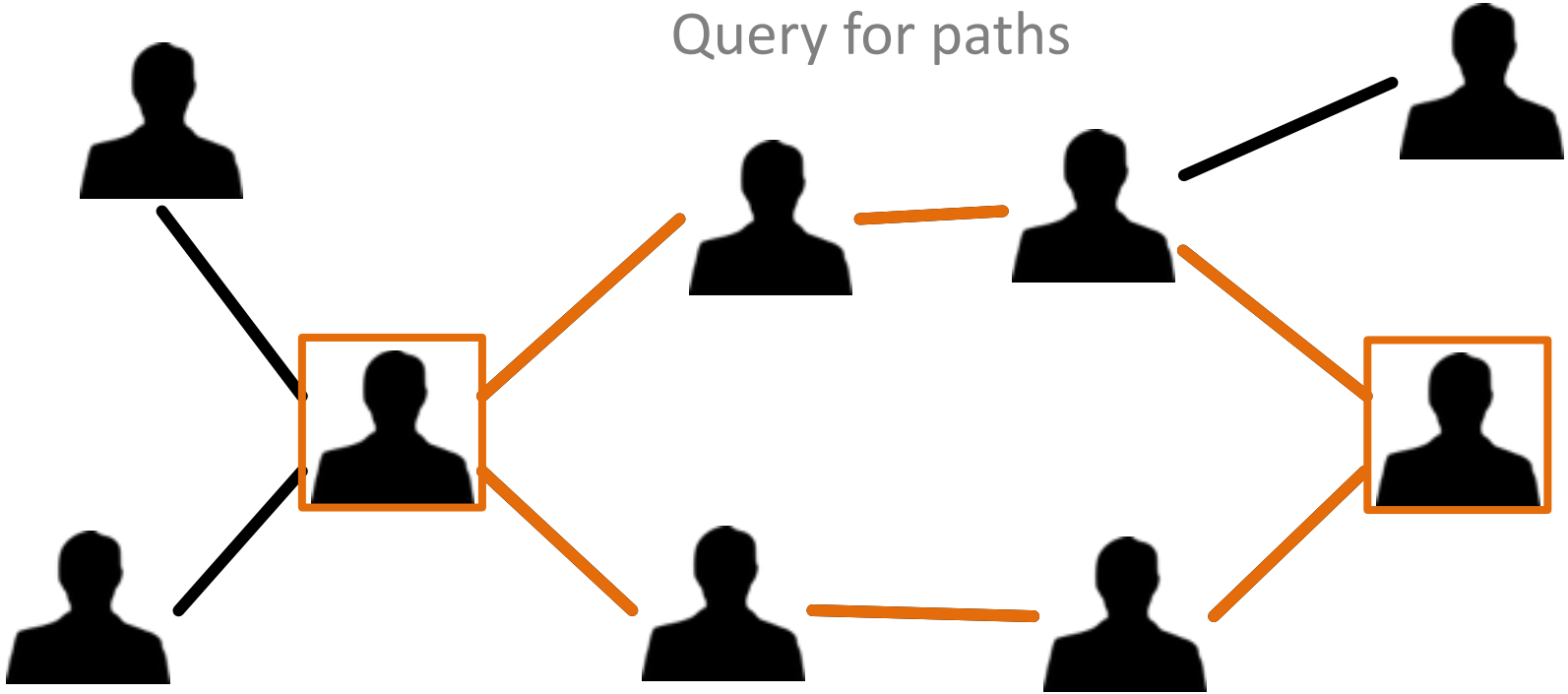
Edges

Attributes

Sets

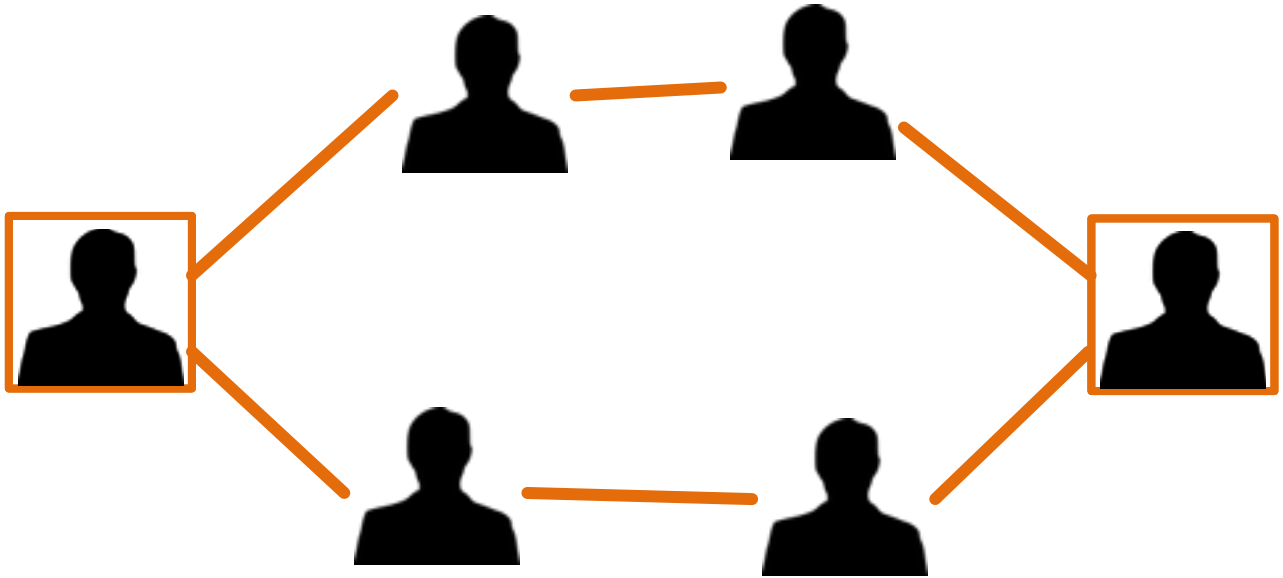
Pathfinder Approach

Query for paths



Pathfinder Approach

Show node-link diagram...

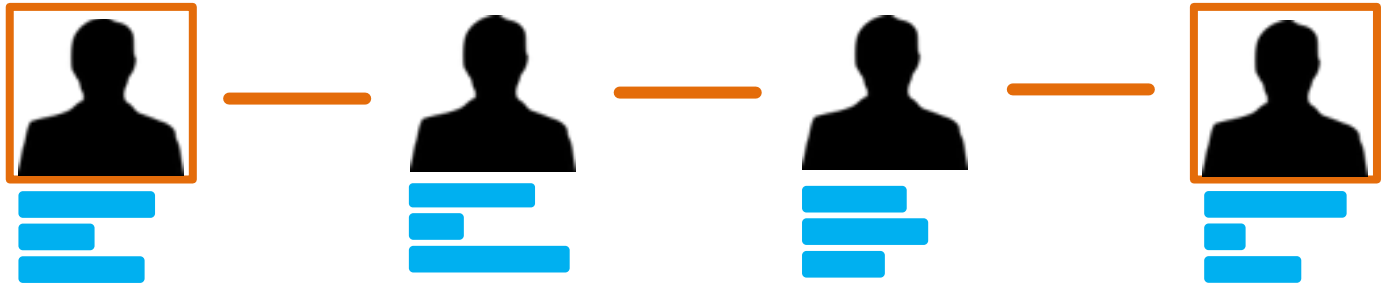


Pathfinder Approach

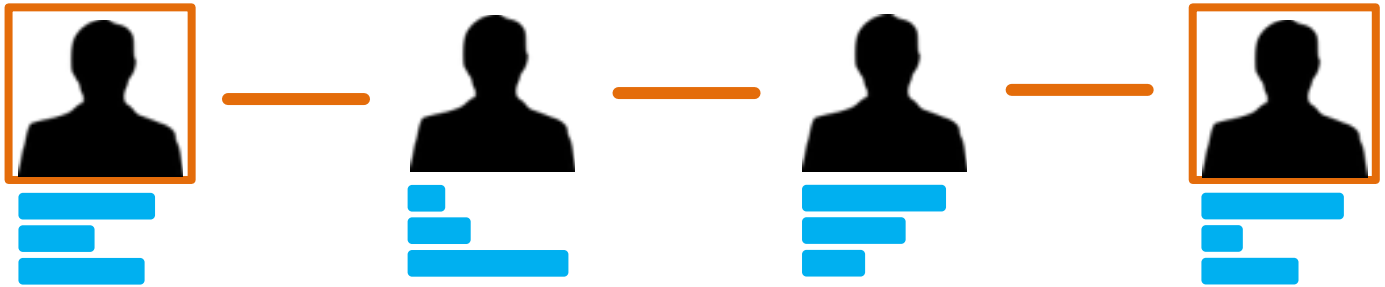
Update ranking to identify important paths

Path Score

1.

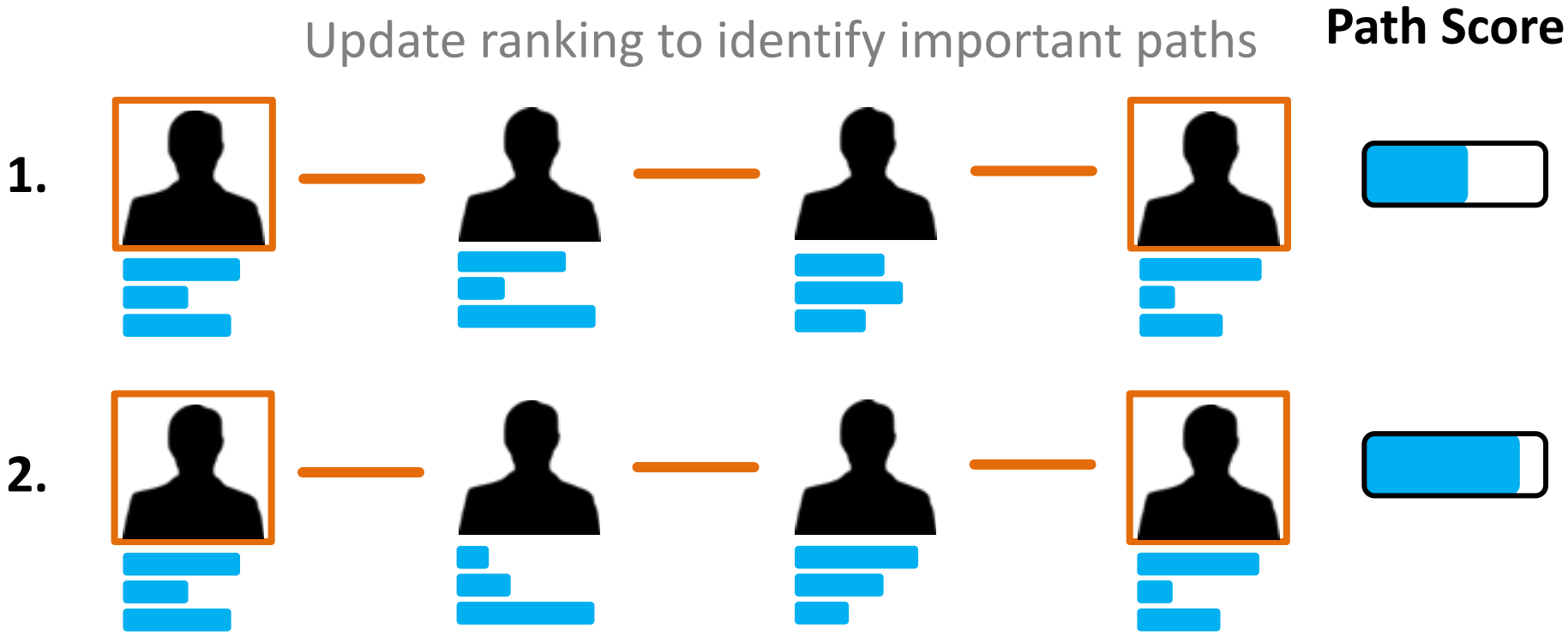


2.

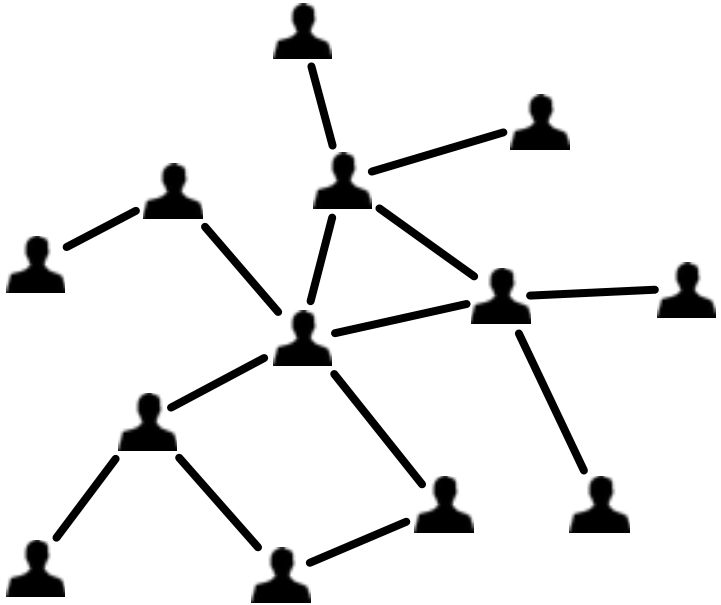


Pathfinder Approach

Update ranking to identify important paths



Coauthor Network



Nodes: Author ***Edges: Co-author-ship***

Generated from TVCG and CHI publication data in DBLP

34k Nodes, 45k Edges

Attributes

Publications

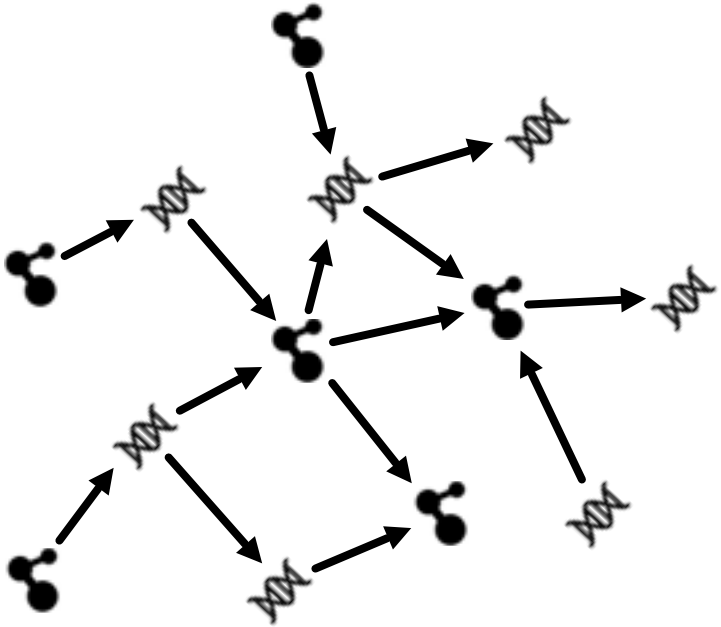
Citations

…

Sets

CHI/TVCG Publications

Biological Network



Generated from

KEGG Pathway Data

11k Nodes, 71k Edges

Gene

Attributes

Gene Expression Data

Copy Number Data

Sets

Biological Pathways

Compound

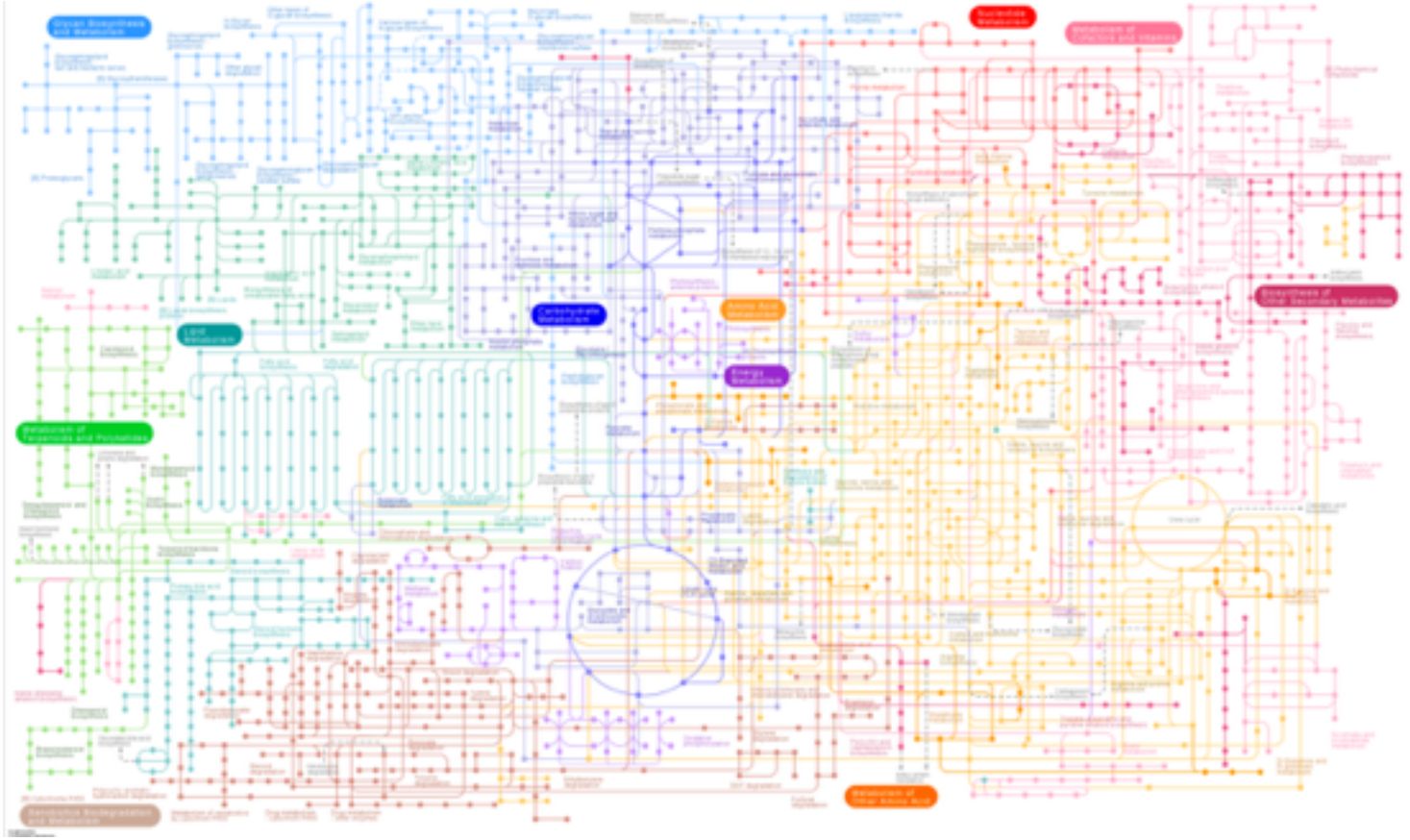
Attributes

-

Sets

Biological Pathways

Example: KEGG Metabolism Overview



Analysis Questions

Two genes are co-expressed. What is their **underlying connection**?

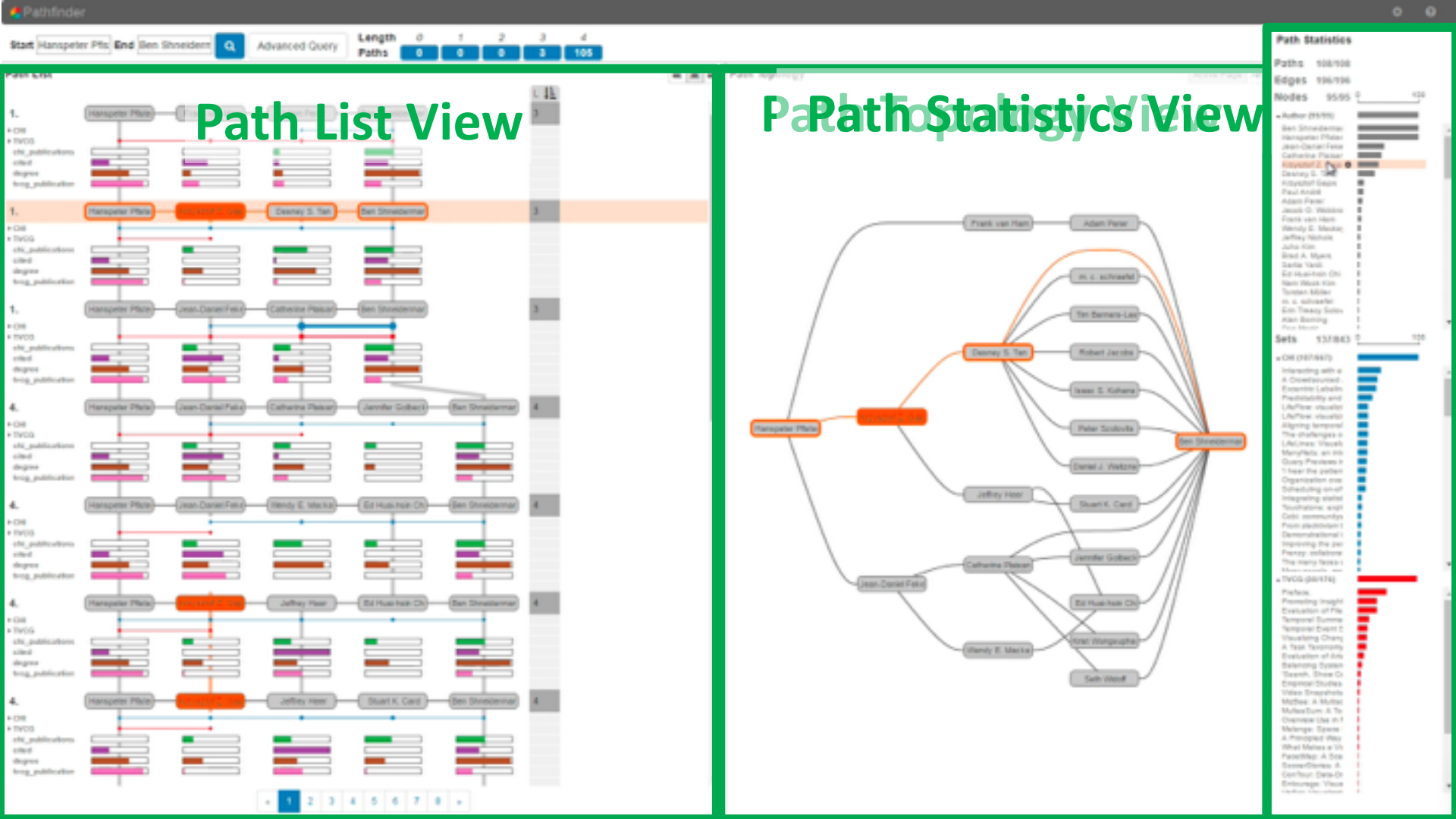
What are **alternative routes** connecting A to B?

How is **gene A connected to Pathway B**?

Is the connection from **A to B the same in many pathways**?

Is the **route connecting A and B active**?

Pathfinder Views


Pathfinder
Start
Hanspeter Pfis
End
Ben Shneiderm
Advanced Query
Length
Paths
0
1
2
3
4
0
0
0
3
105
Path List View
Path Topology View
Path Statistics View
Path Statistics
Paths
Edges
Nodes
Sets
Hanspeter Pfister
Frank van Ham
Adam Perer
Danney S. Tan
Robert Jacobs
Isaac S. Kohane
Peter Szolovits
Daniel J. Weitzne
Ben Shneiderman
Jeffrey Heer
Stuart K. Card
Catherine Plaisan
Jennifer Golbeck
Jean-Daniel Fekz
Ed Huai-hsin Chi
Wendy E. Mackay

Path Query and Search

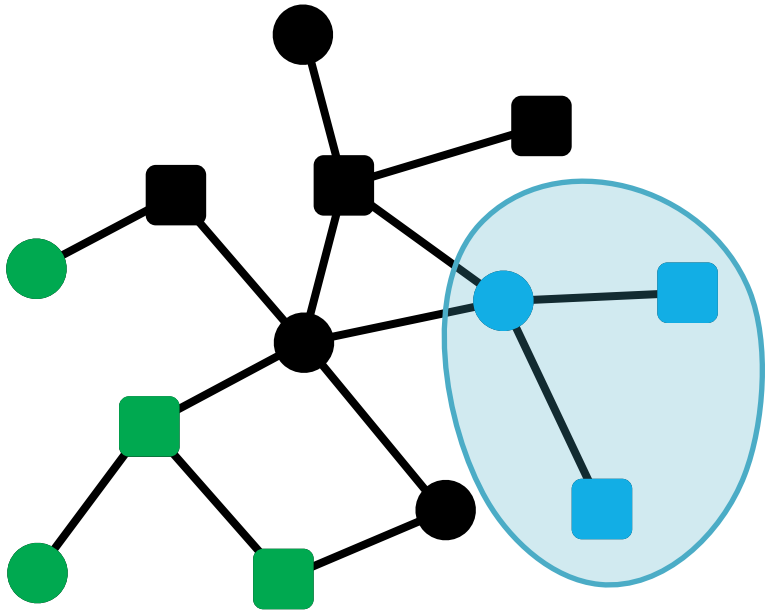
Finding Path Relationships

Query and Search

Specify **start** and **end**

Start/end can be node lists

Start/end can be defined through set membership

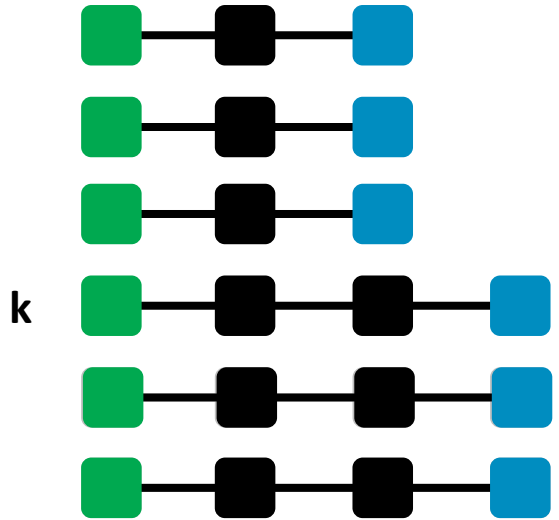


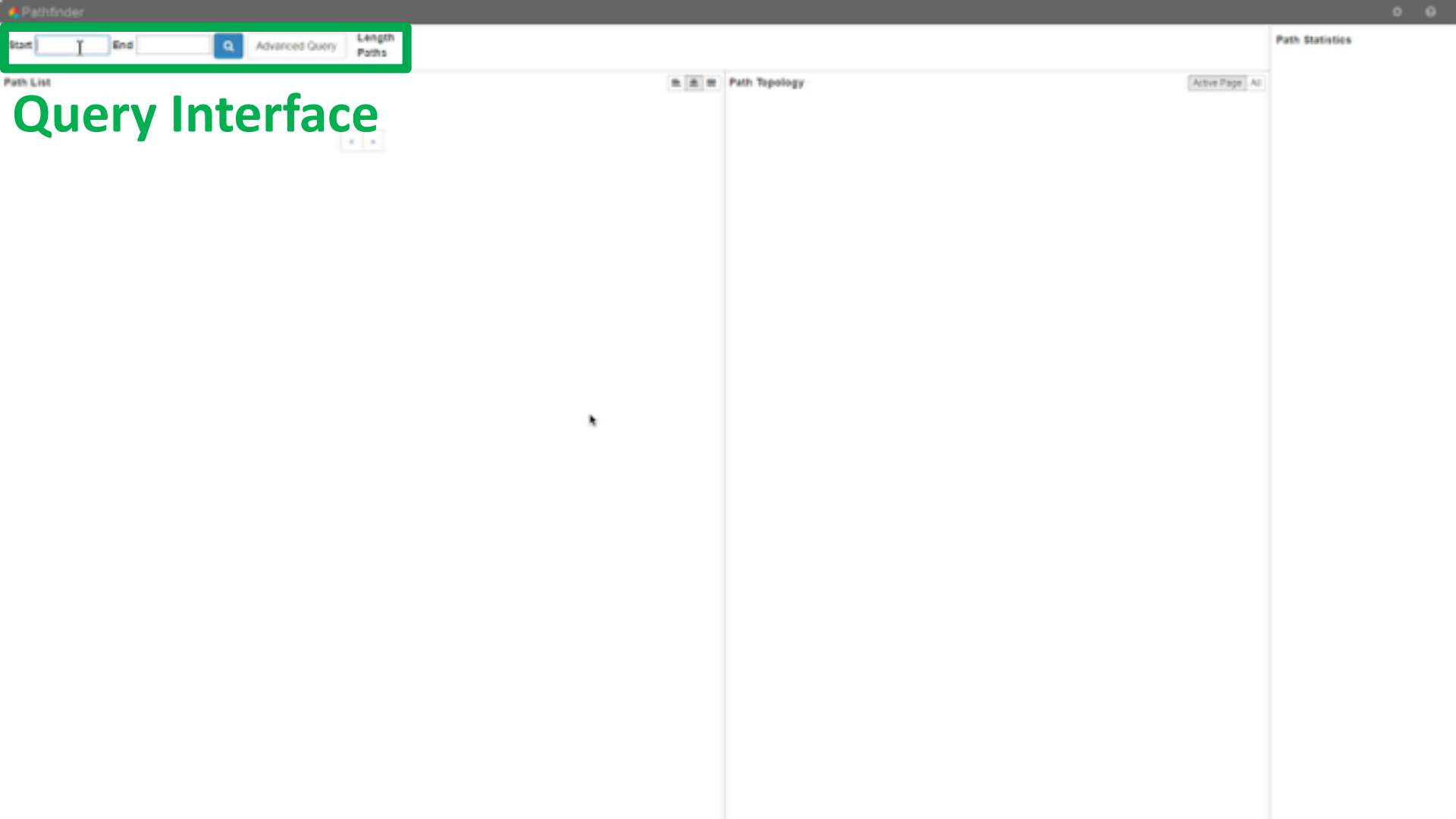
Query and Search

Specify **start** and **end**

K-shortest path search

Continue until all path of length(k) are found

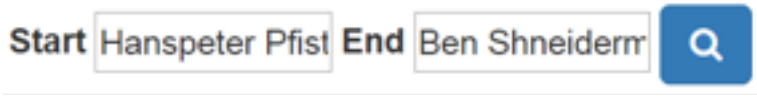




Query Interface

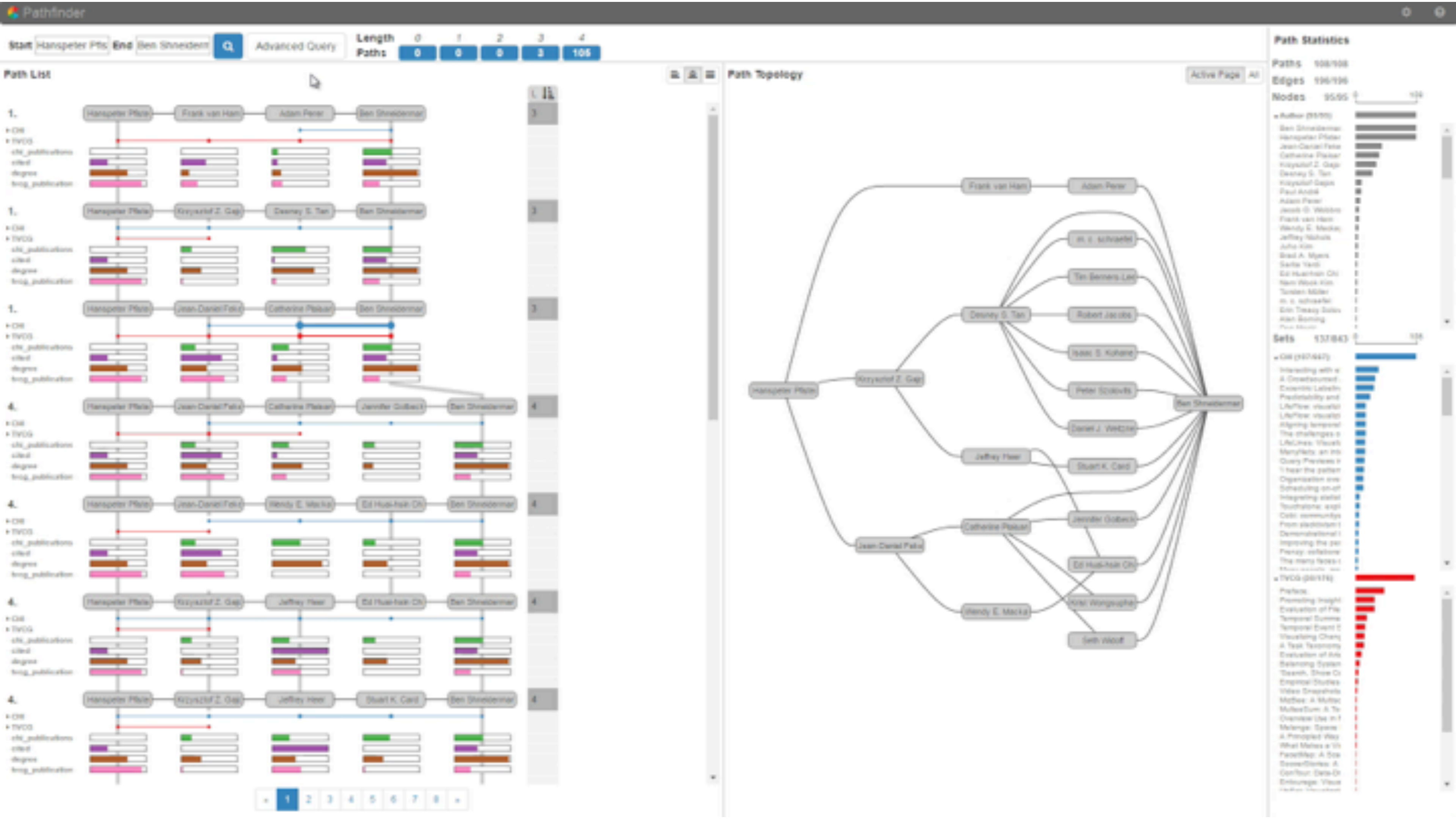
Query Interface

Simple



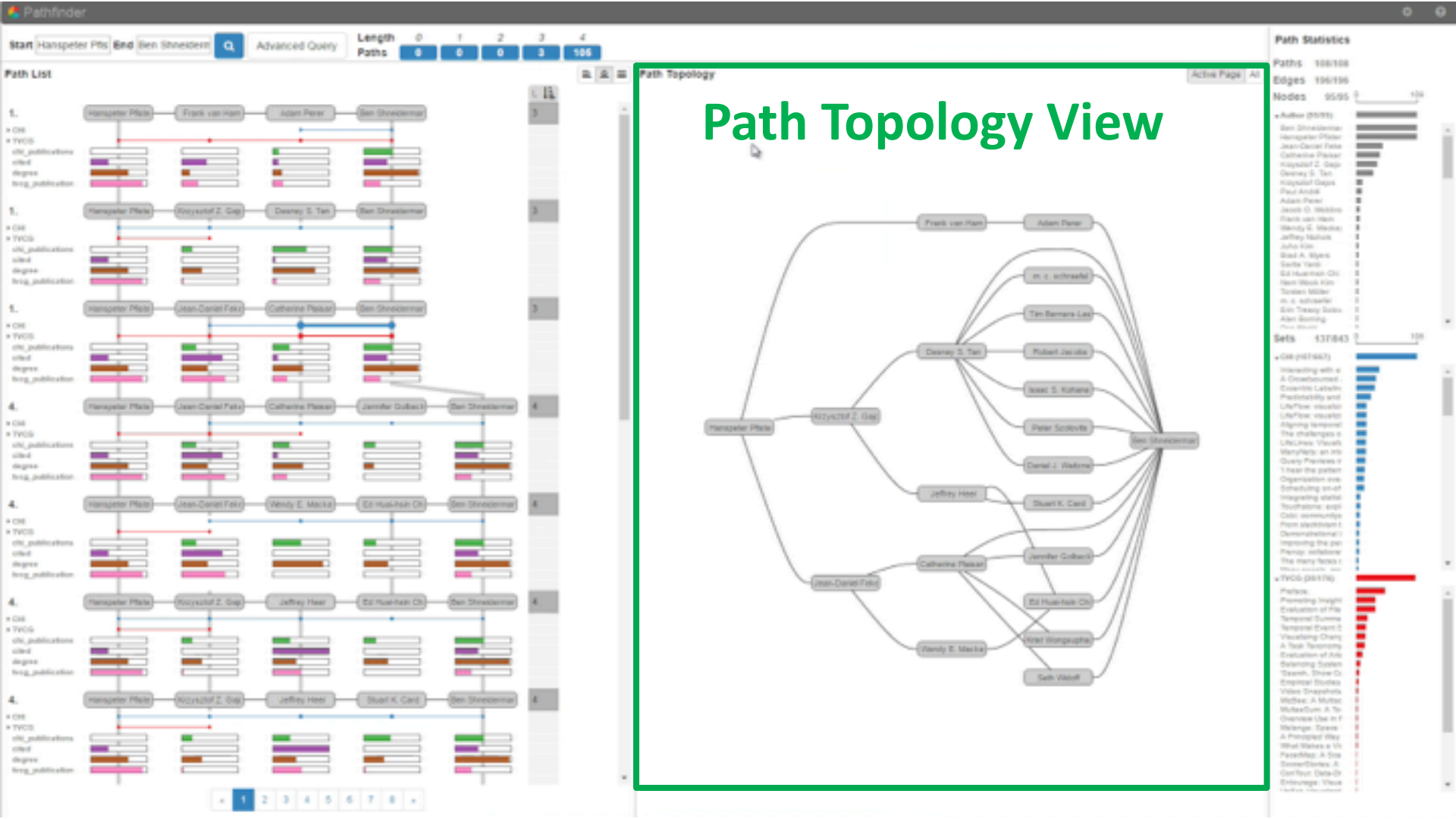
Length	0	1	2	3	4
Paths	0	0	0	3	105

Advanced


Pathfinder
Start
End
Advanced Query
Length
Paths
0
1
2
3
4
105
Path List
Path Topology
Active Page
All
Path Statistics
Paths
Edges
Nodes
Sets
Hanspeter Pfister
Frank van Ham
Adam Perer
Ben Shneiderman
Krzysztof Z. Gajos
Desney S. Tan
Jean-Daniel Fekete
Catherine Plaisant
Jennifer Golbeck
Wendy E. Mackay
Ed Huai-hsin Chi
Jeffrey Heer
Stuart K. Card
Tim Berners Lee
Robert Jacobs
Isaac S. Kohane
Daniel J. Wetzel
Seth Woolf
CHI
TVCG
cited
degree
1
2
3
4
5
6
7
8

Path Topology View

Getting an Overview

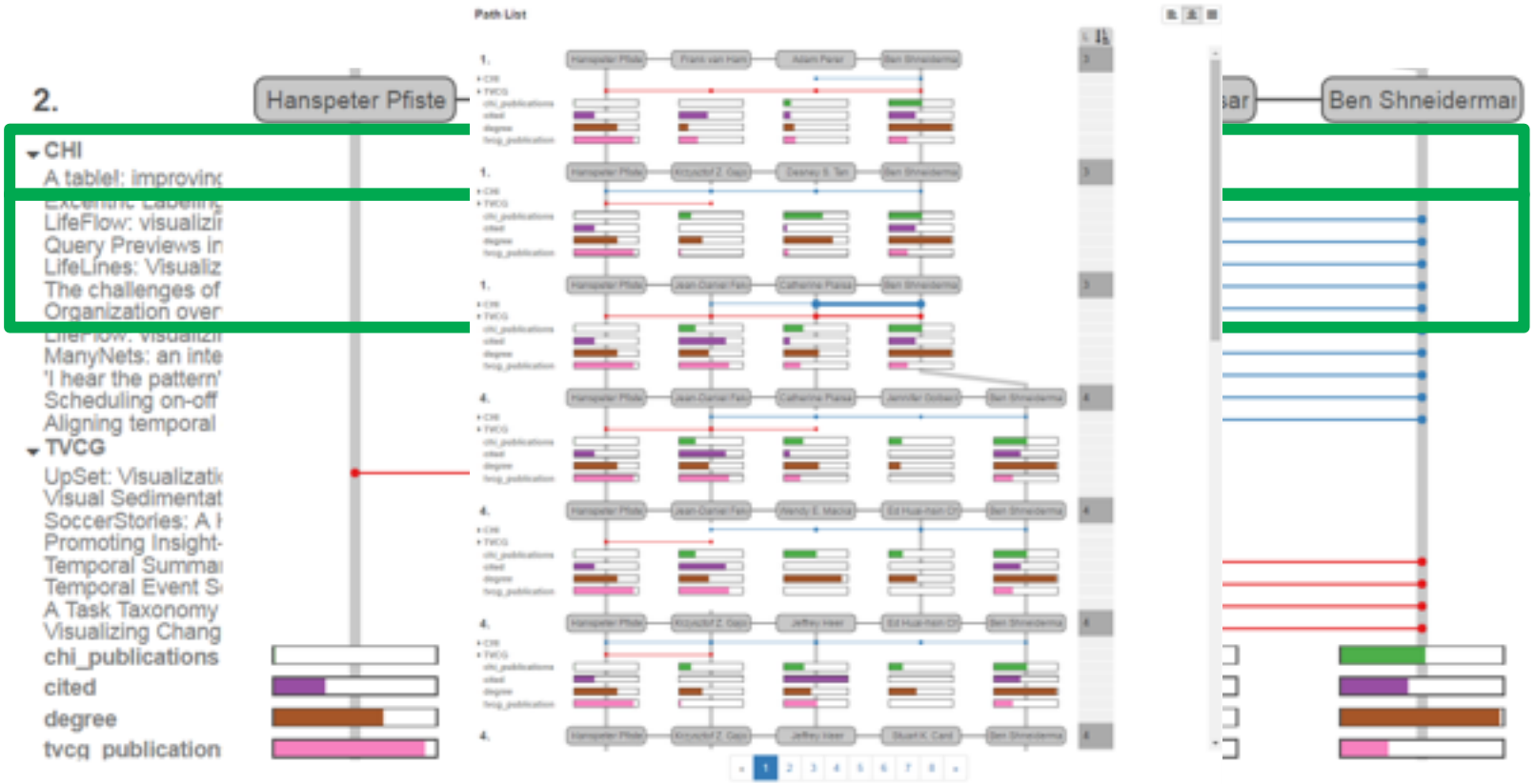


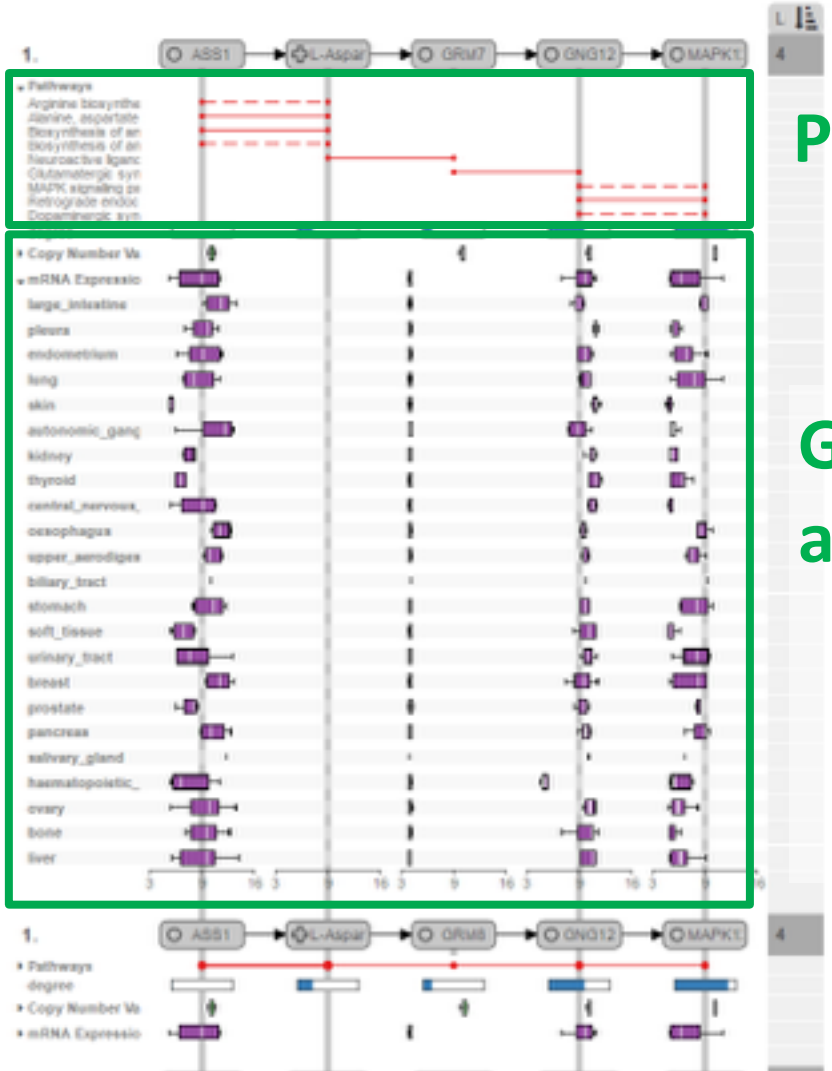
Path Topology View

Path List View

Investigating Paths in Detail

Path Representation



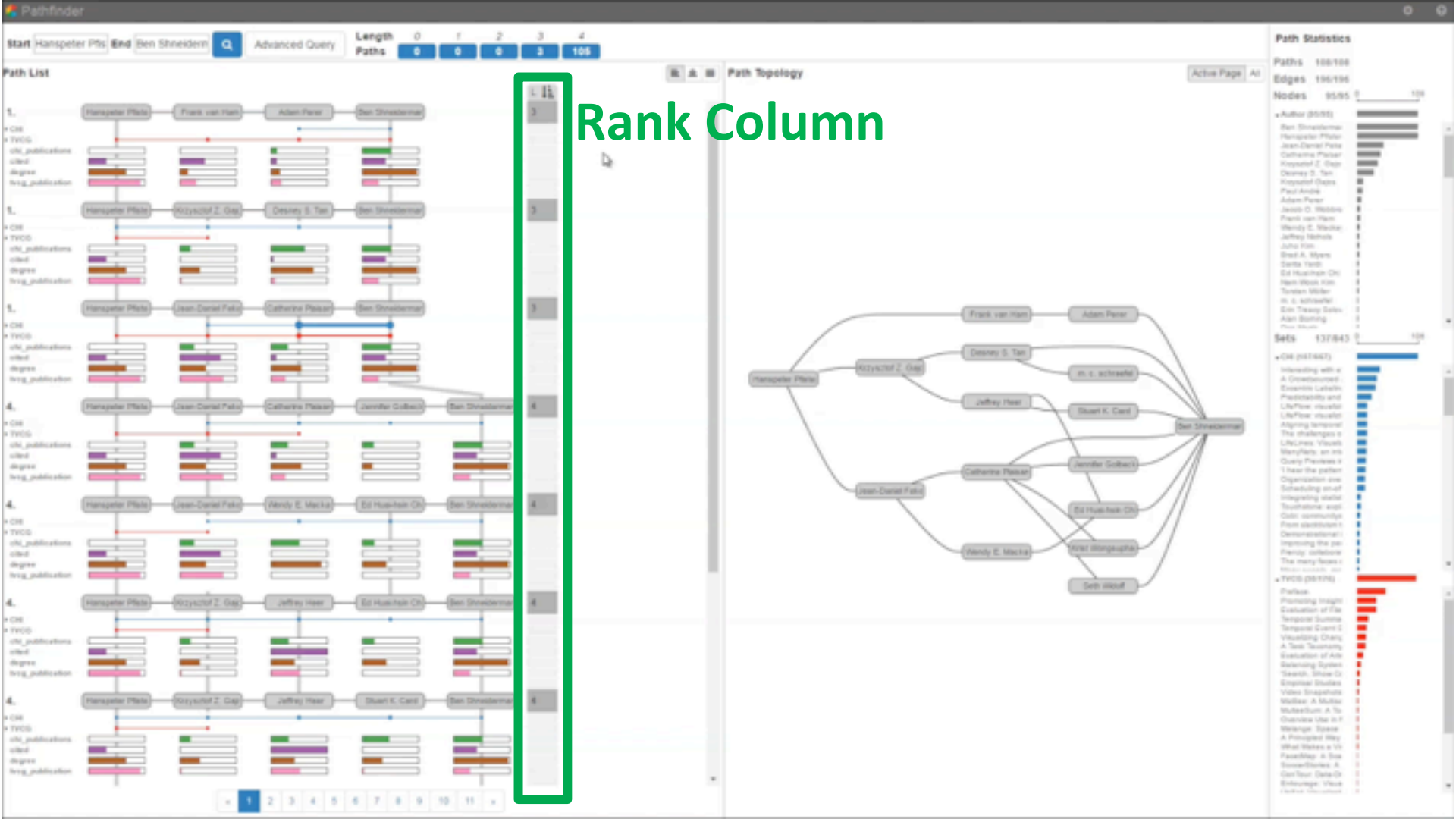


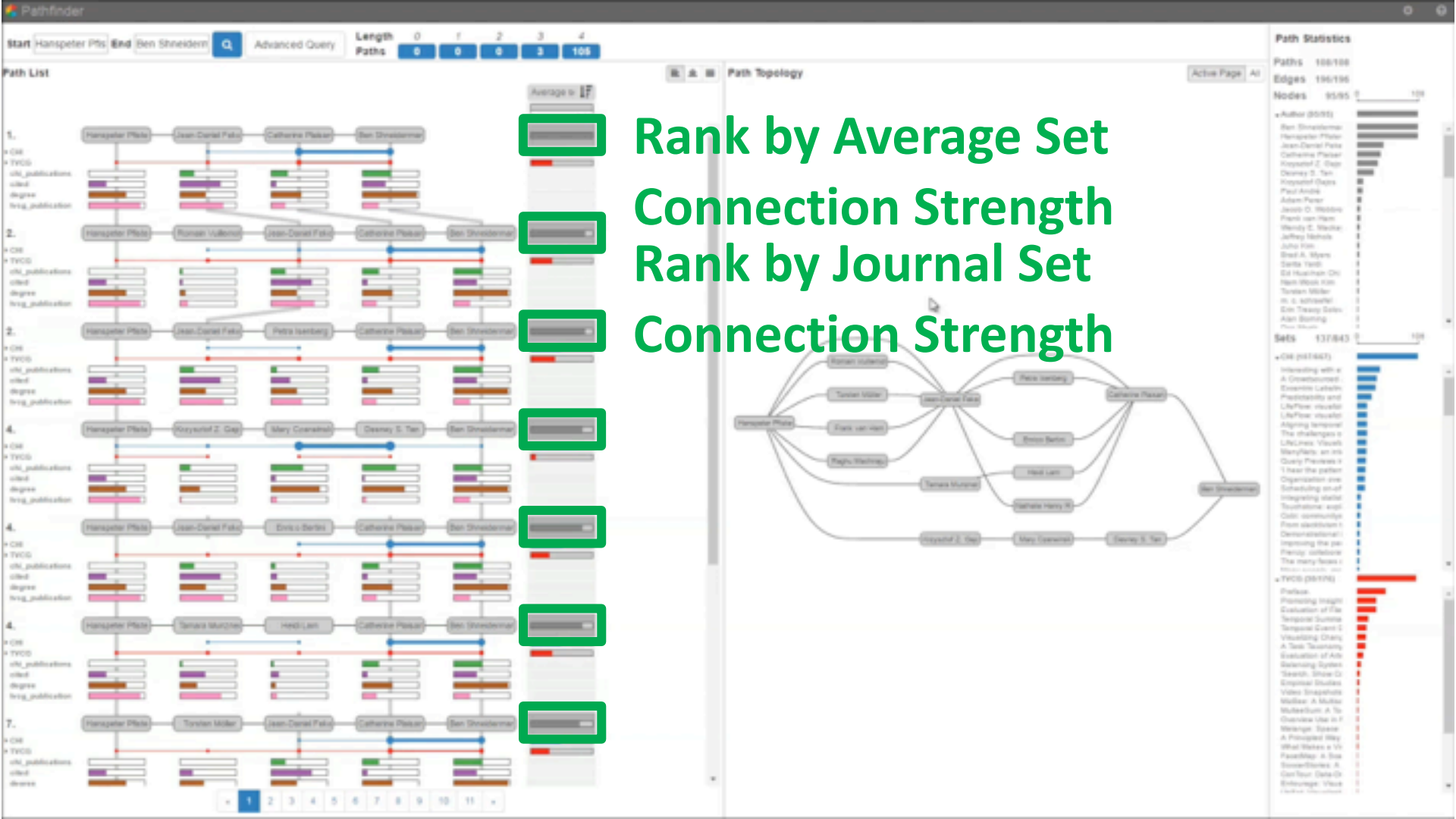
Pathways

Grouped Copy Number and Gene Expression Data

Path Ranking

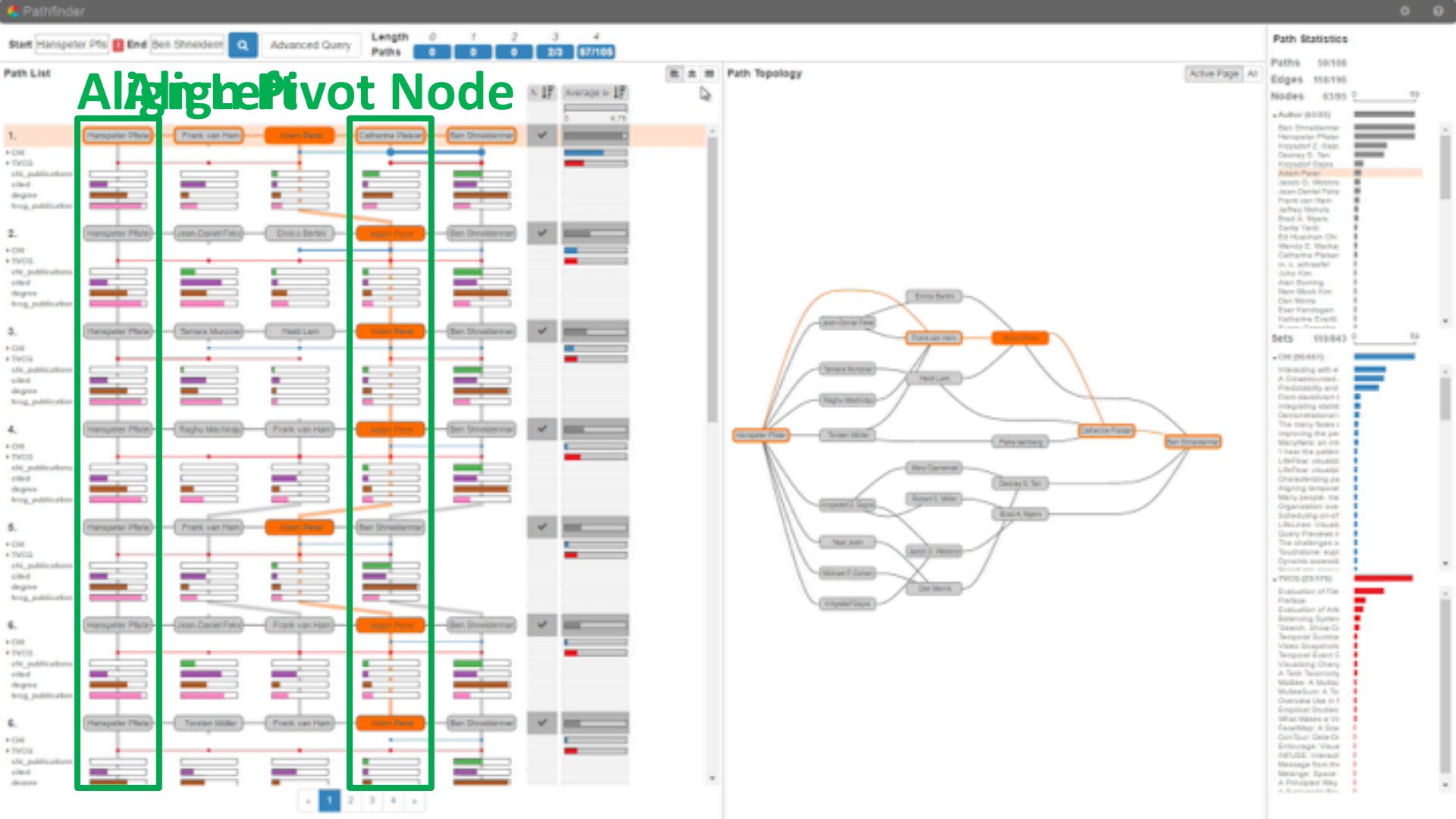
Identifying Relevant Paths

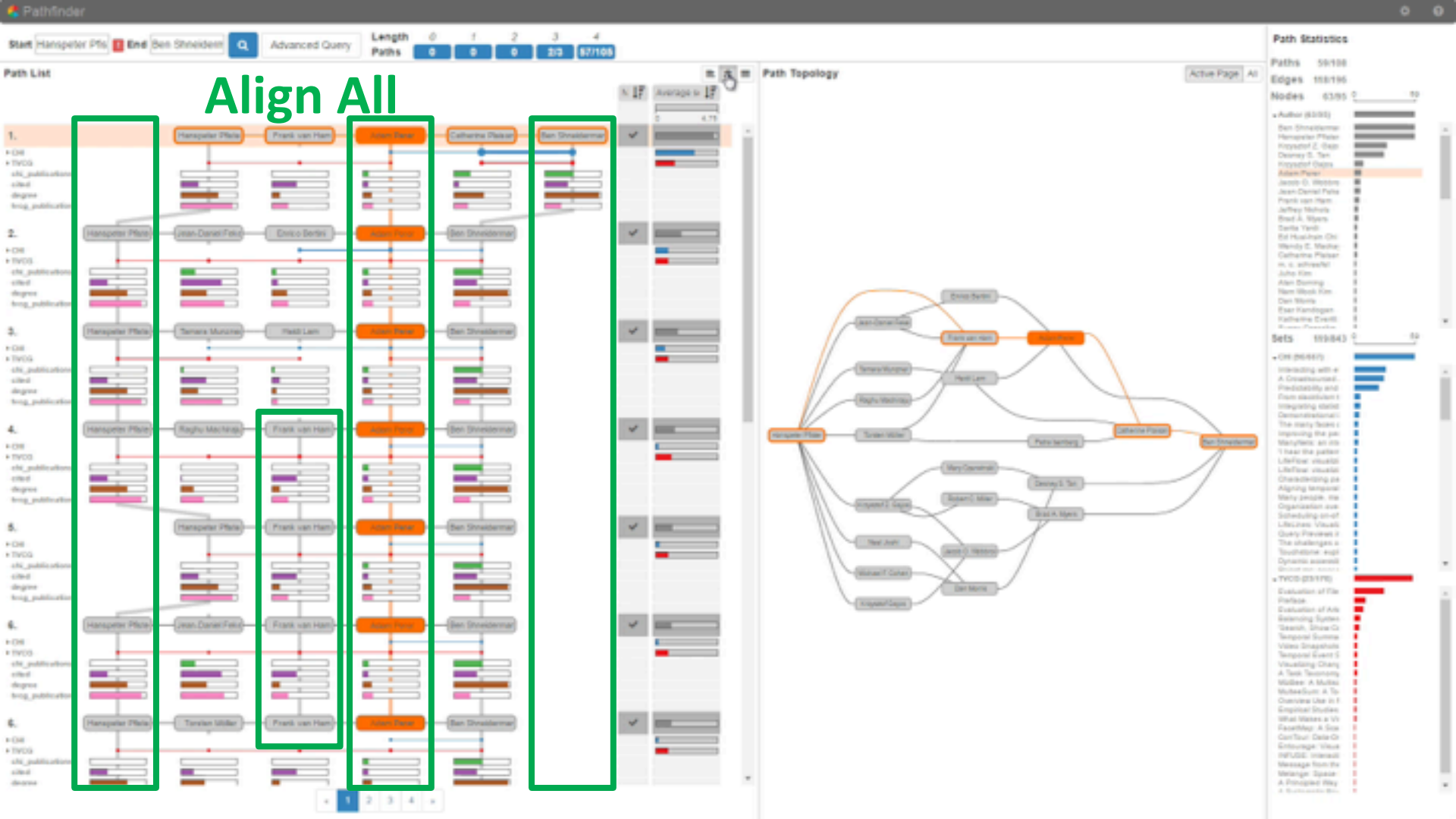

Pathfinder
Start
End
Advanced Query
Length
Paths
Path List
Path Topology
Path Statistics
Rank Column


Pathfinder
Start
End
Advanced Query
Length
Paths
Path List
Path Topology
Path Statistics
Rank by Average Set Connection Strength
Rank by Journal Set Connection Strength

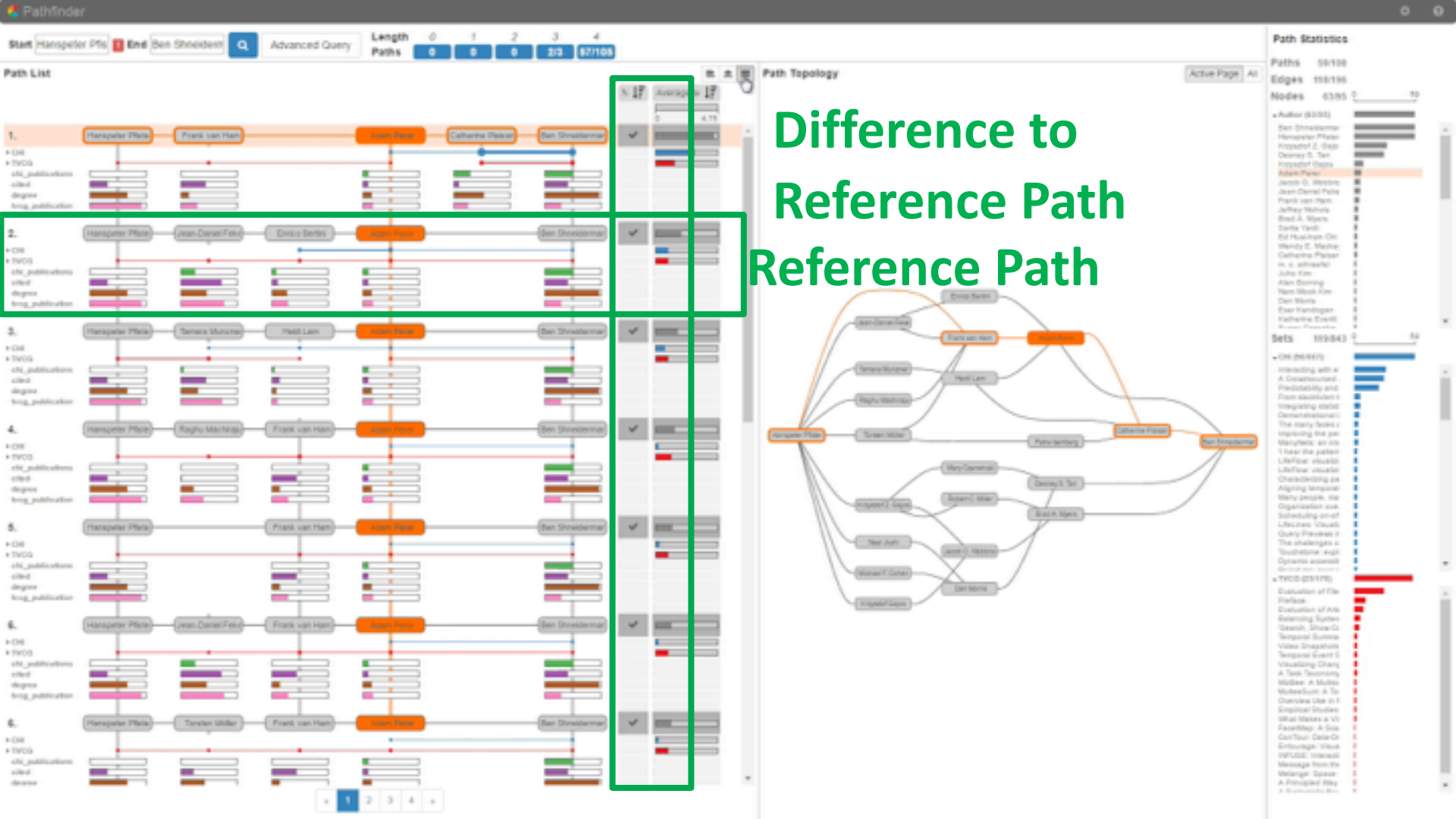
Path Alignment

Facilitating Path Comparison


Pathfinder
Start
End
Advanced Query
Length
Paths
Path List
Aligned Pivot Node
Path Topology
Path Statistics
Paths
Edges
Nodes
Sets

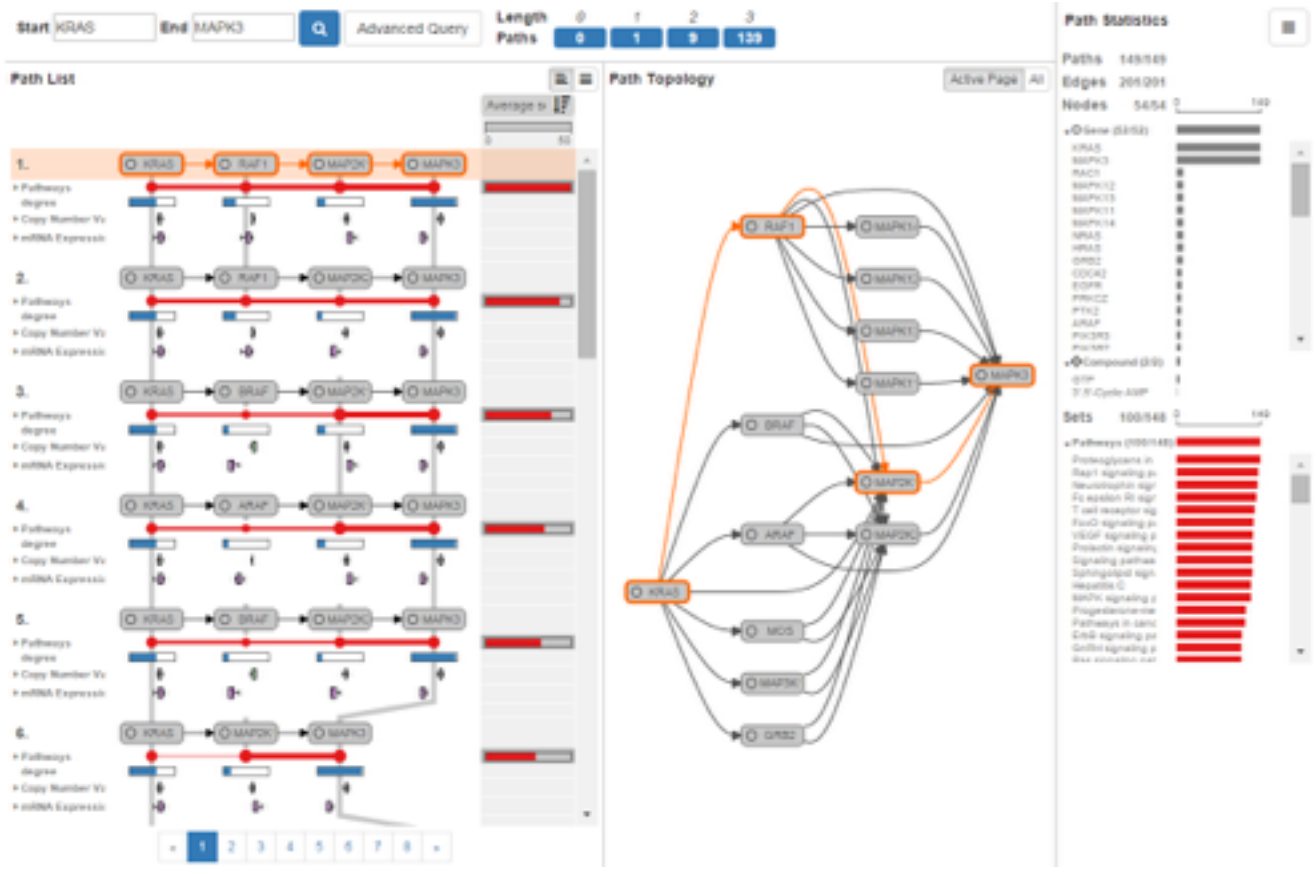

Pathfinder
Start
End
Advanced Query
Length
Paths
Path List
Align All
Path Topology
Active Page
Path Statistics
Paths
Edges
Nodes
Sets

Comparison to Reference Path


Pathfinder
Start
Hanspeter Pfis
End
Ben Shneiderm
Advanced Query
Length
Paths
0
1
2
3
4
0
0
0
2/2
57/109
Path List
Path Topology
Active Page
All
Path Statistics
Paths
Edges
Nodes
Sets
Difference to
Reference Path
Reference Path

Case Study:
Biological Network

ERK-MAPK signaling cascade



Search Paths from KRAS to MAPK3

Ranking by set connection reveals ERK-MAPK signaling cascade

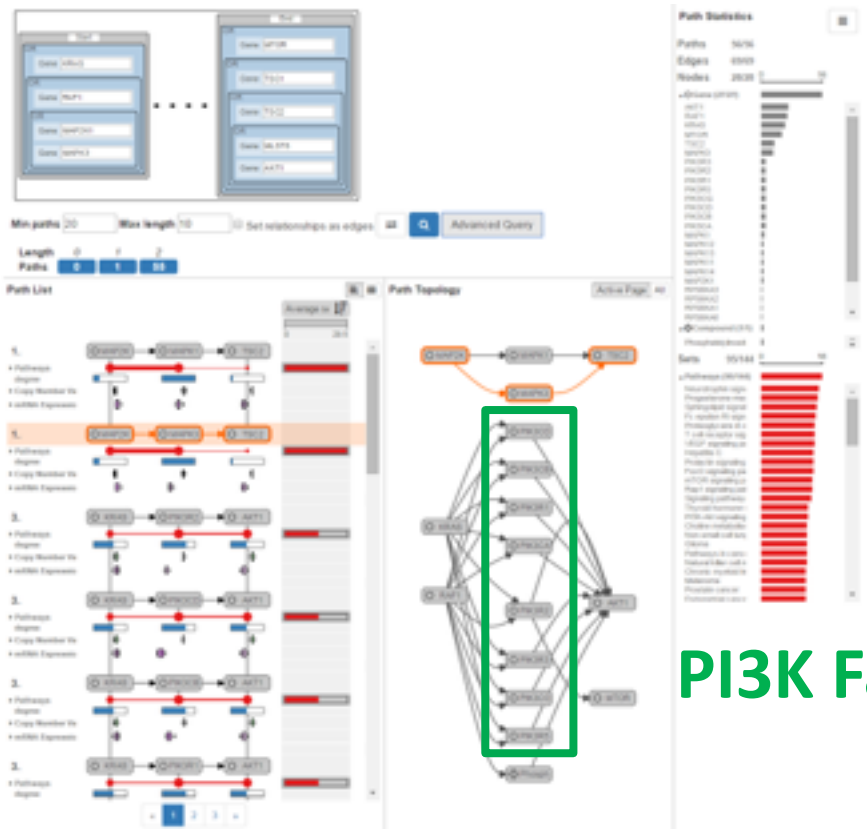
ERK-MAPK signaling cascade

This cascade is important in many pathways.

ERK-MAPK signaling cascade

This cascade is expressed ubiquitously in all cancer tissue types

Crosstalk between ERK-MAPK and mTOR



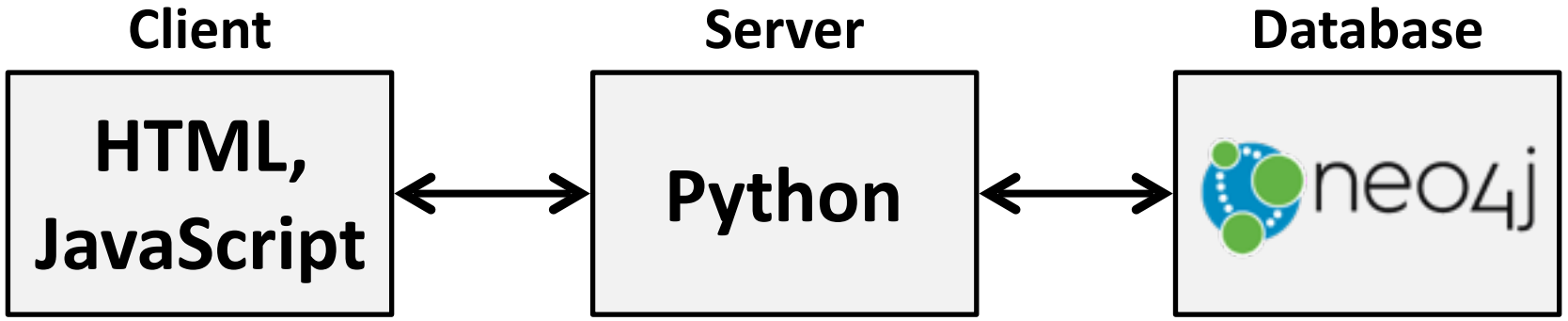
Complex query between various players in both cascades

Reveals possible crosstalk via MAPK3 and crosstalk via P13K Family

PI3K Family

Implementation

Implementation

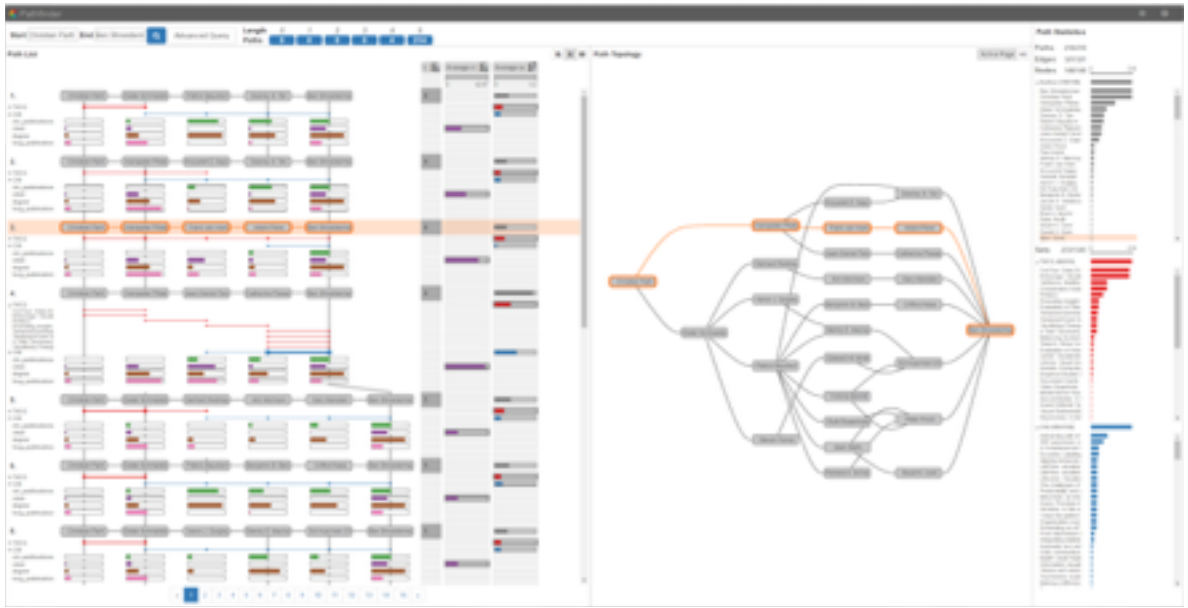


Immediate feedback to queries through intermediate results

Rendering: D3, Dagre

CALEYDO

Pathfinder in a Nutshell



Query-based Technique for Path Analysis

Results in Topology- and List View

Path Exploration by

- Query Refinement
- Ranking

Questions?

CALEYDO

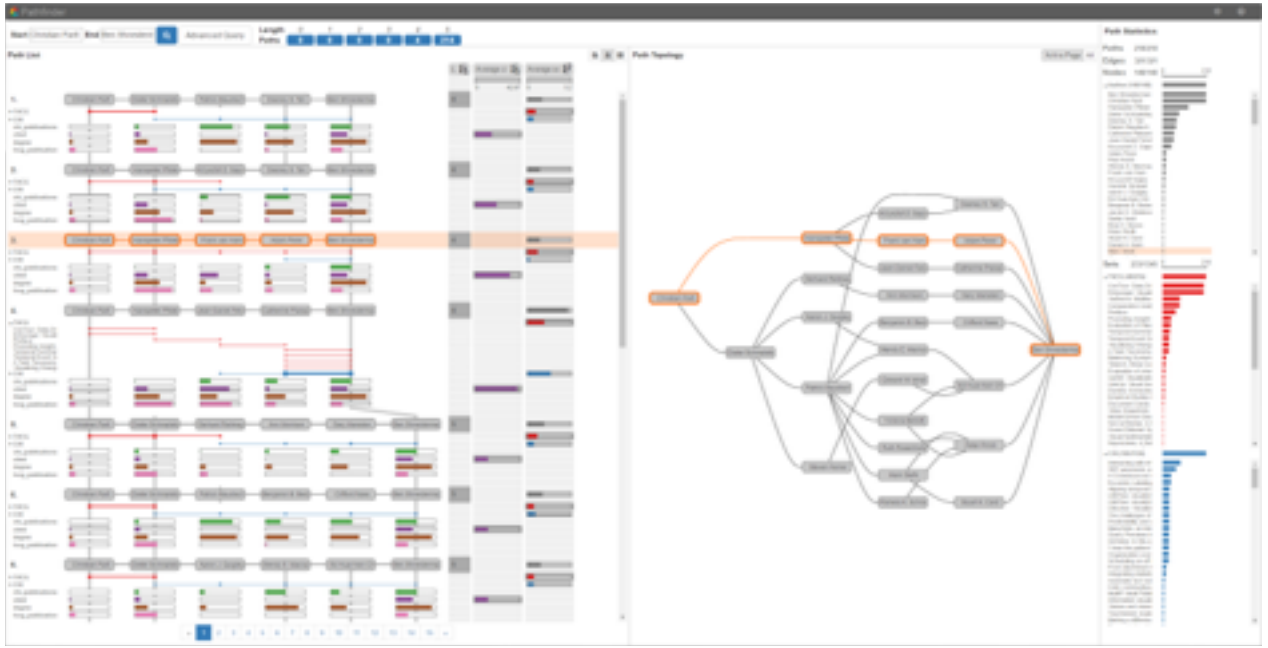
pathfinder.caleydo.org

Interactive Demo:

demo.caleydo.org/pathfinder/

Source Code:

github.com/Caleydo/pathfinder/



Alexander Lex

alex@sci.utah.edu
http://alexander-lex.net
http://vdl.sci.utah.edu

TU Graz Graz University of Technology
Pfizer
JKU Johannes Kepler Universität Linz
Harvard School of Engineering and Applied Sciences
visualization design lab
The University of Utah

We're hiring PostDocs and accept PhD Students!

visualization design lab

Miriah Meyer

Alexander Lex

http://goo.gl/Zmm4L8

Future Work

Integrate with general graph exploration tools

Provide overview of similar paths

- Grouping based on similarity
- Dimensionality reduction and show in scatterplot

Cost-based path search

Scalability

Network Size

Coauthor Network: 34k nodes, 45k edges

Biological Network: 11k nodes, 71k edges

600+ Paths

Path length

Depends on screen size/resolution and node size

7-10 hops, more with scrolling