

Visualizing set-typed data

Chen He

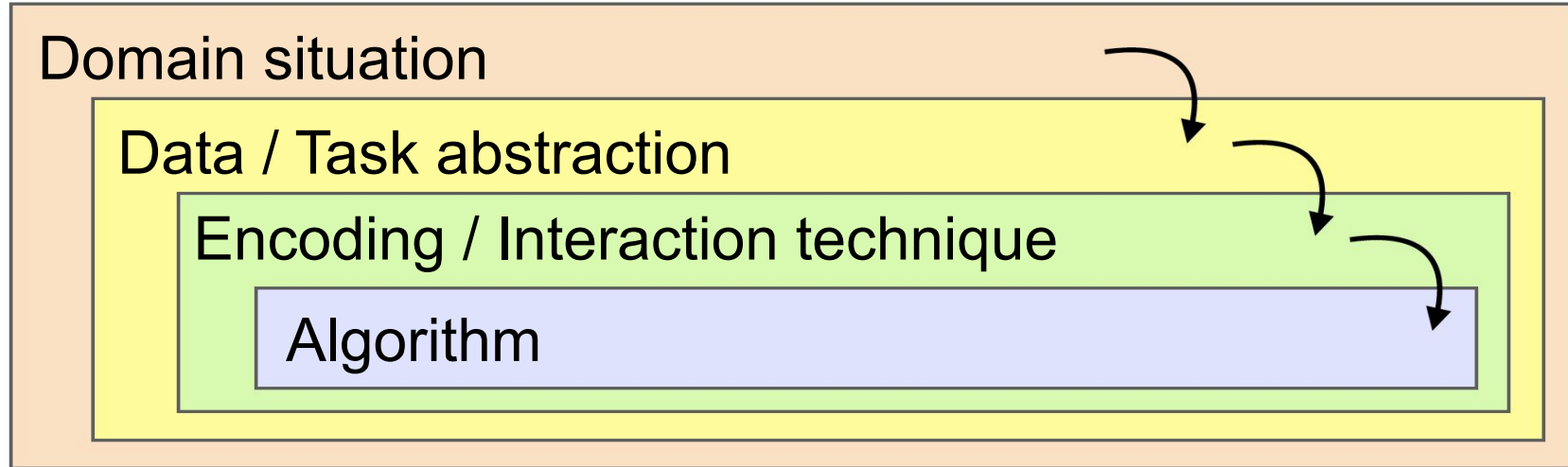
Slides acknowledgements: **Luana Micallef**, and **Miriah Meyer** (Univ. Utah)

What are set-typed data?

Data items are often grouped into sets based on specific properties.



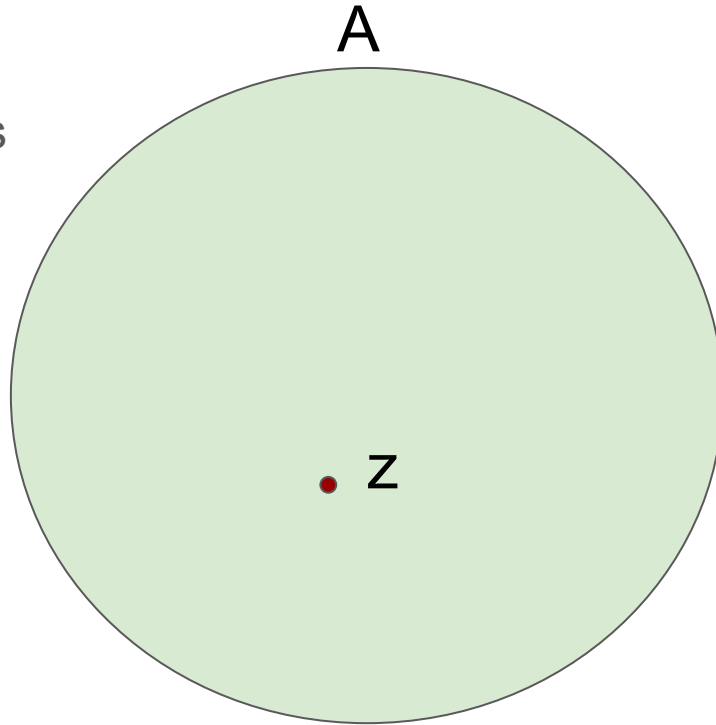
Data abstraction: Set-typed data



Set theory

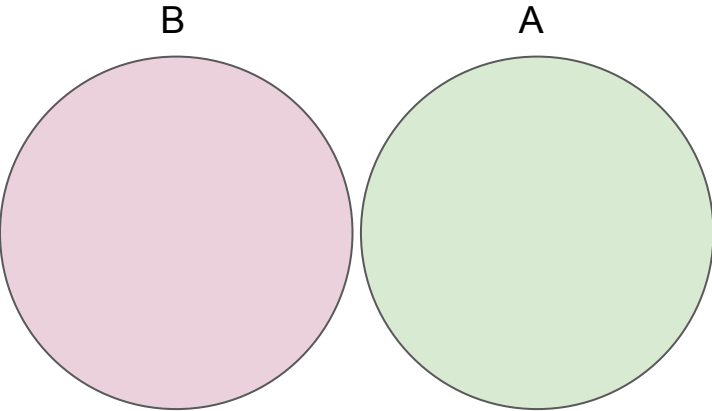
Set: A collection of elements

Element $z \in$ set A

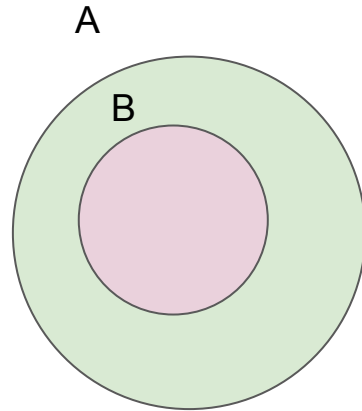


Set theory: Set relations

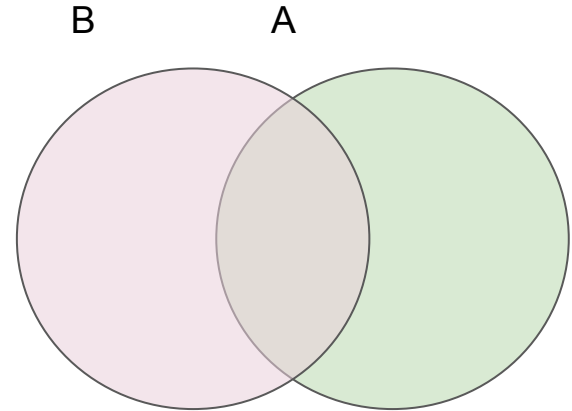
Exclusion



Containment

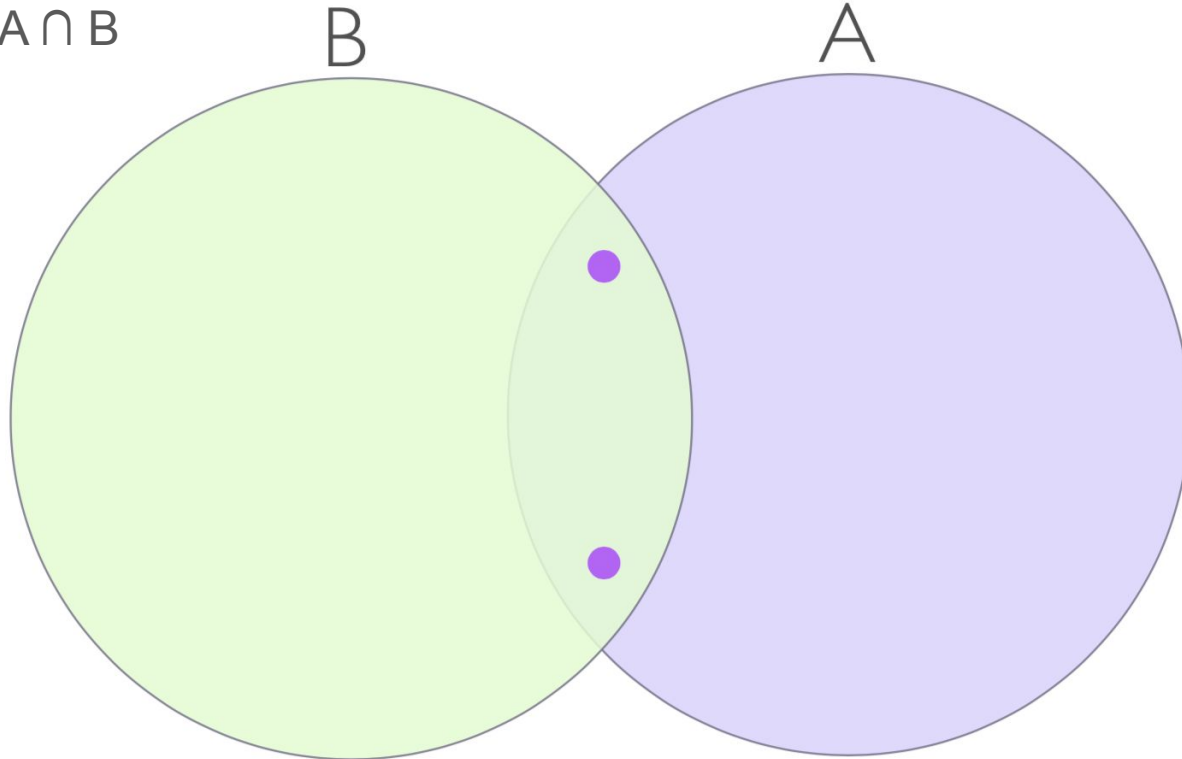


Intersection



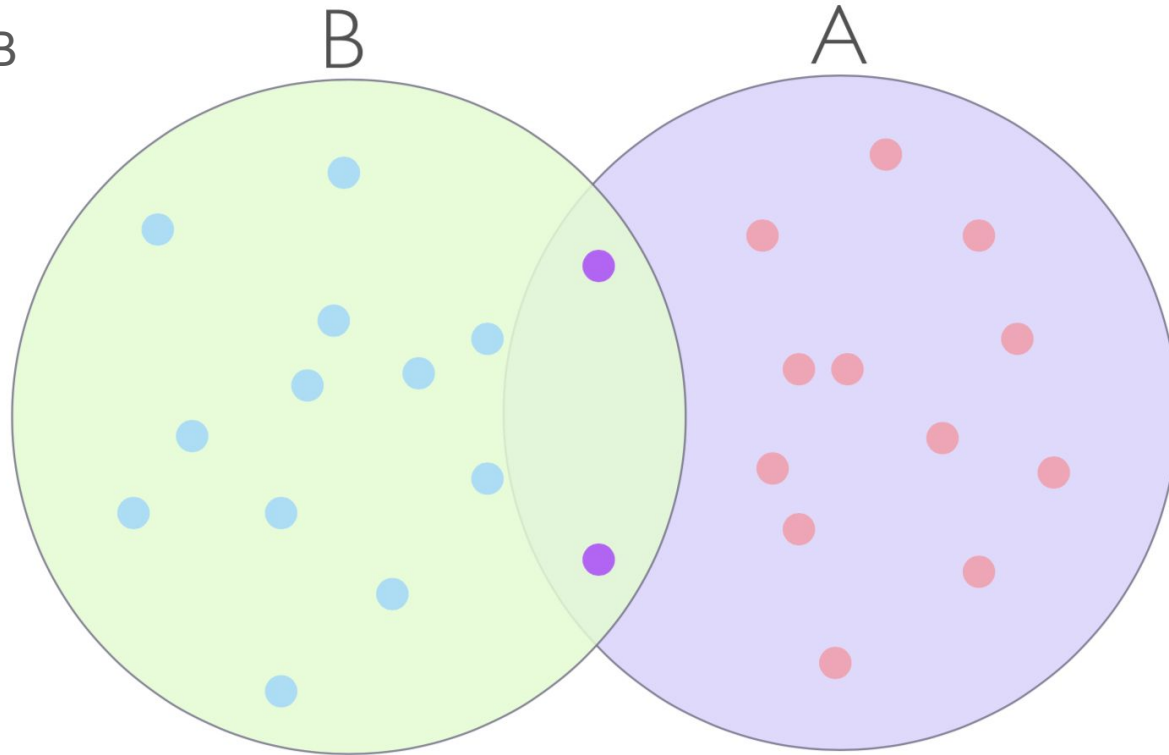
Set theory: Set operations

Intersection: $A \cap B$



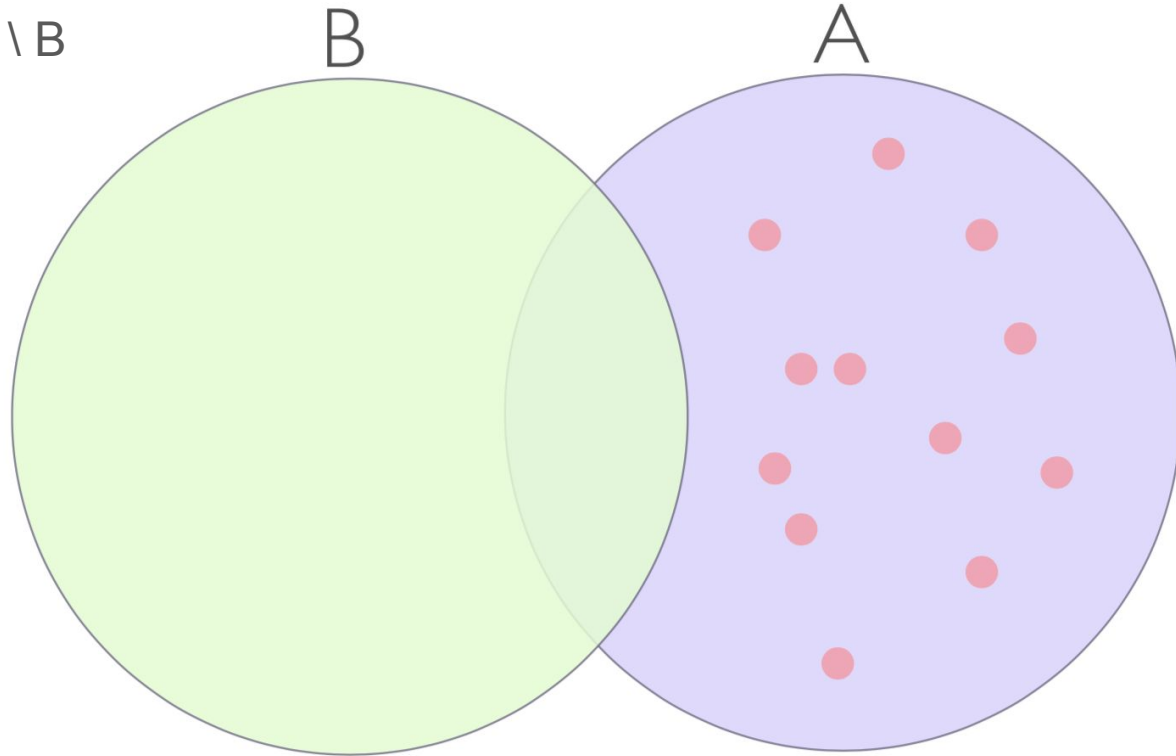
Set theory: Set operations

Union: $A \cup B$



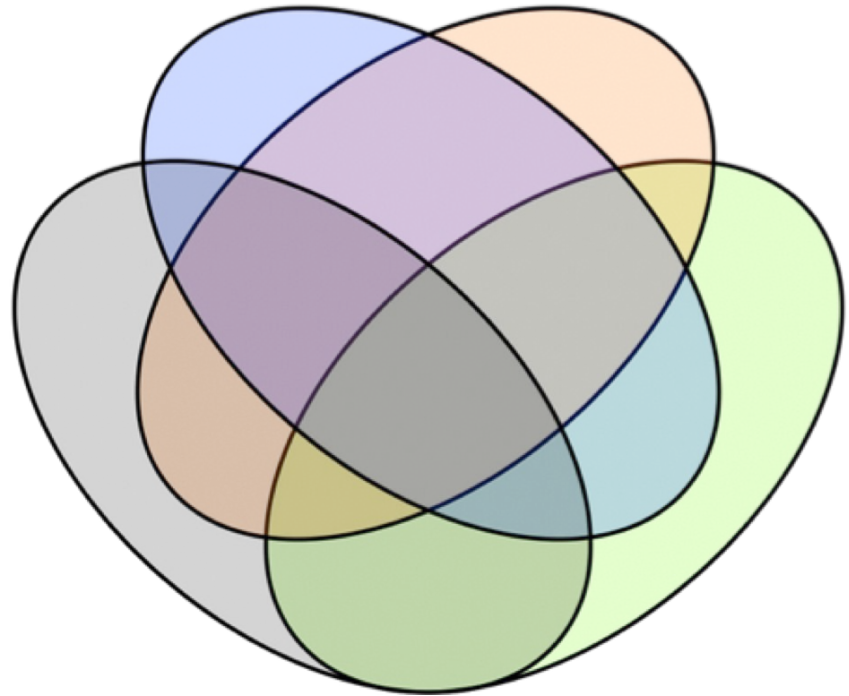
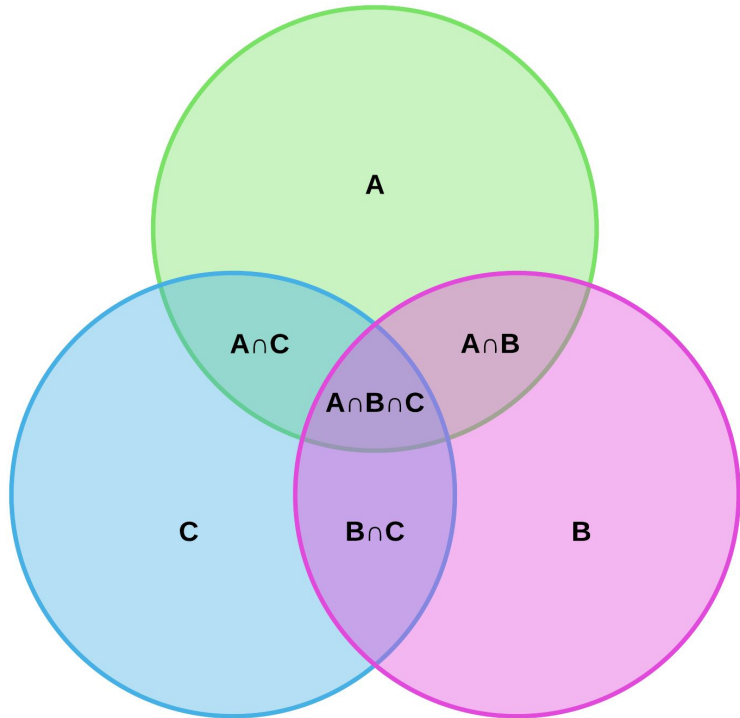
Set theory: Set operations

Difference: $A \setminus B$



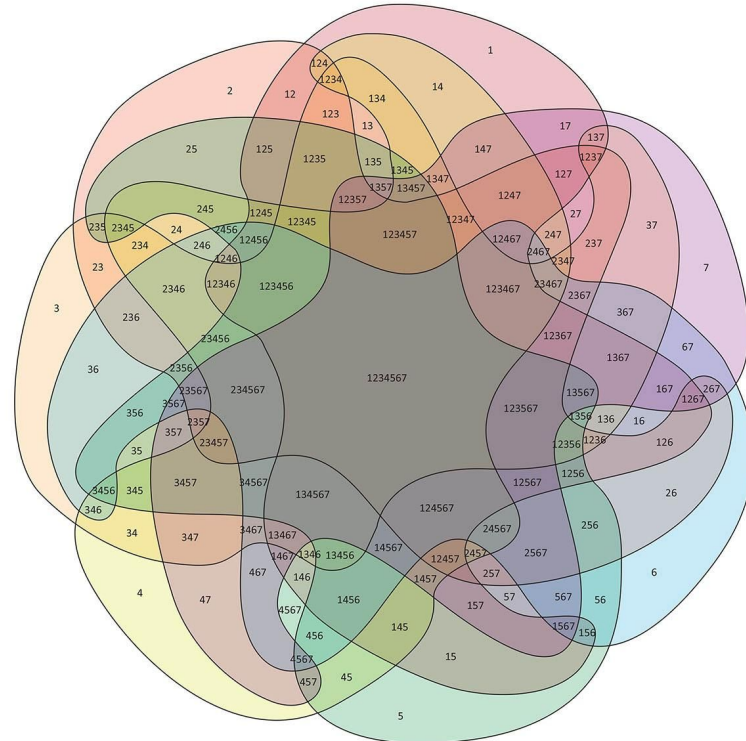
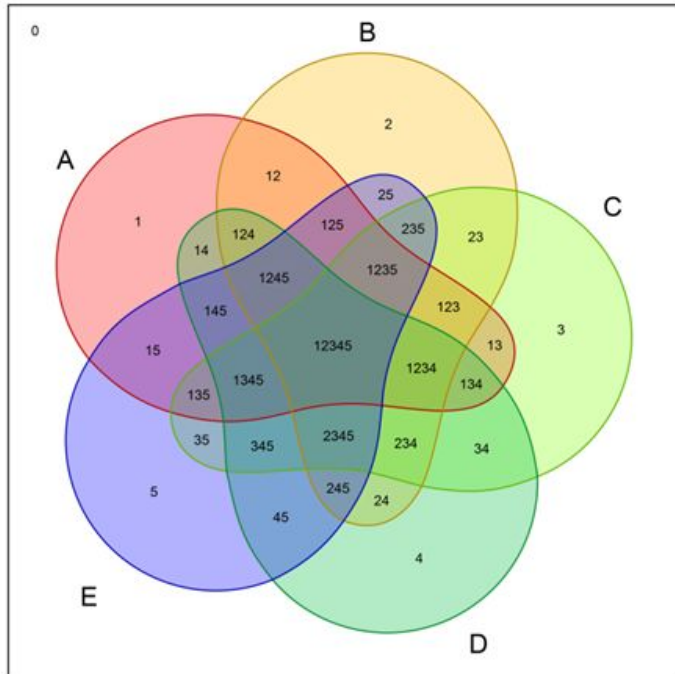
Venn diagrams

Show all possible set relations.



Venn diagrams

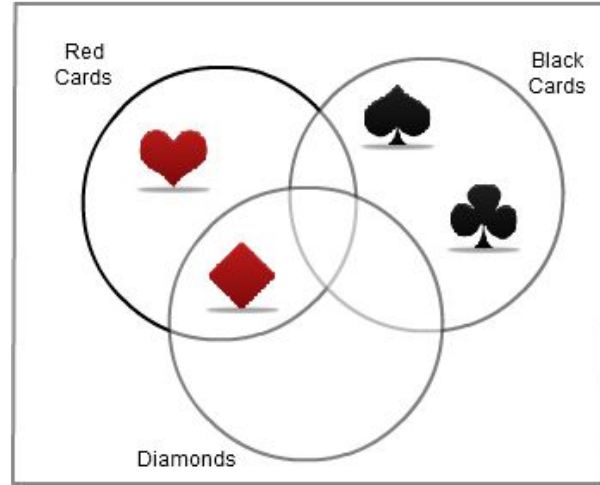
Get messy fast.



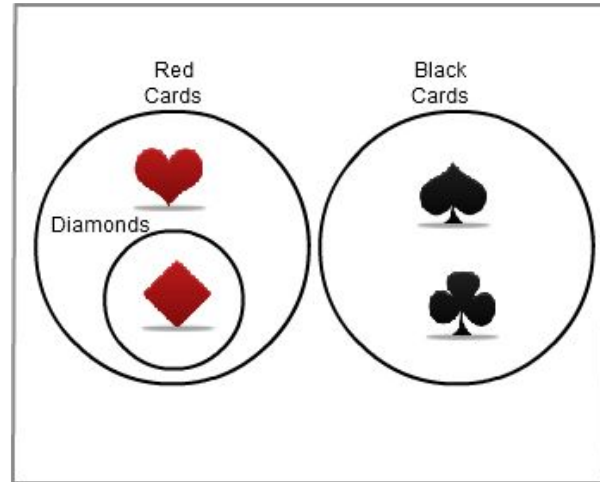
Euler diagrams

Only show existing set relations.

V
E
N
N



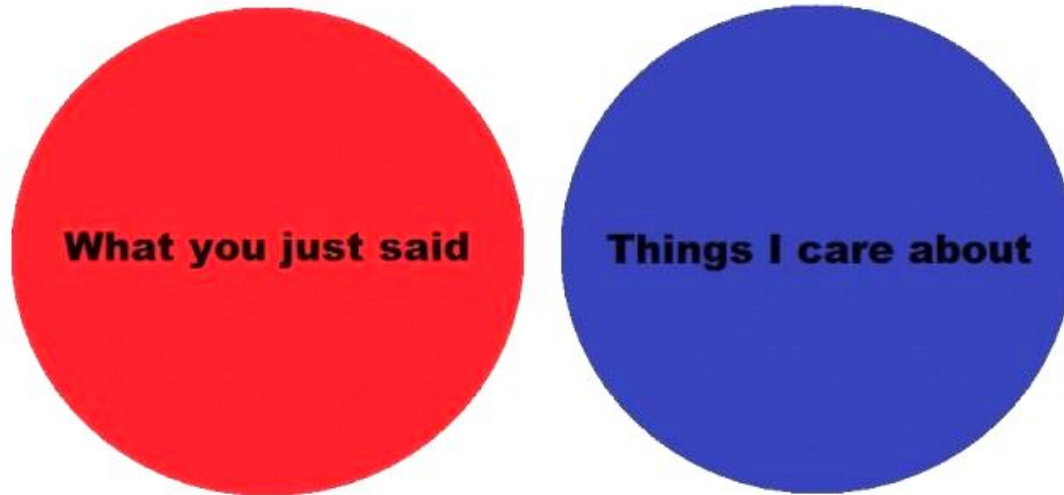
E
U
L
E
R



Euler diagrams

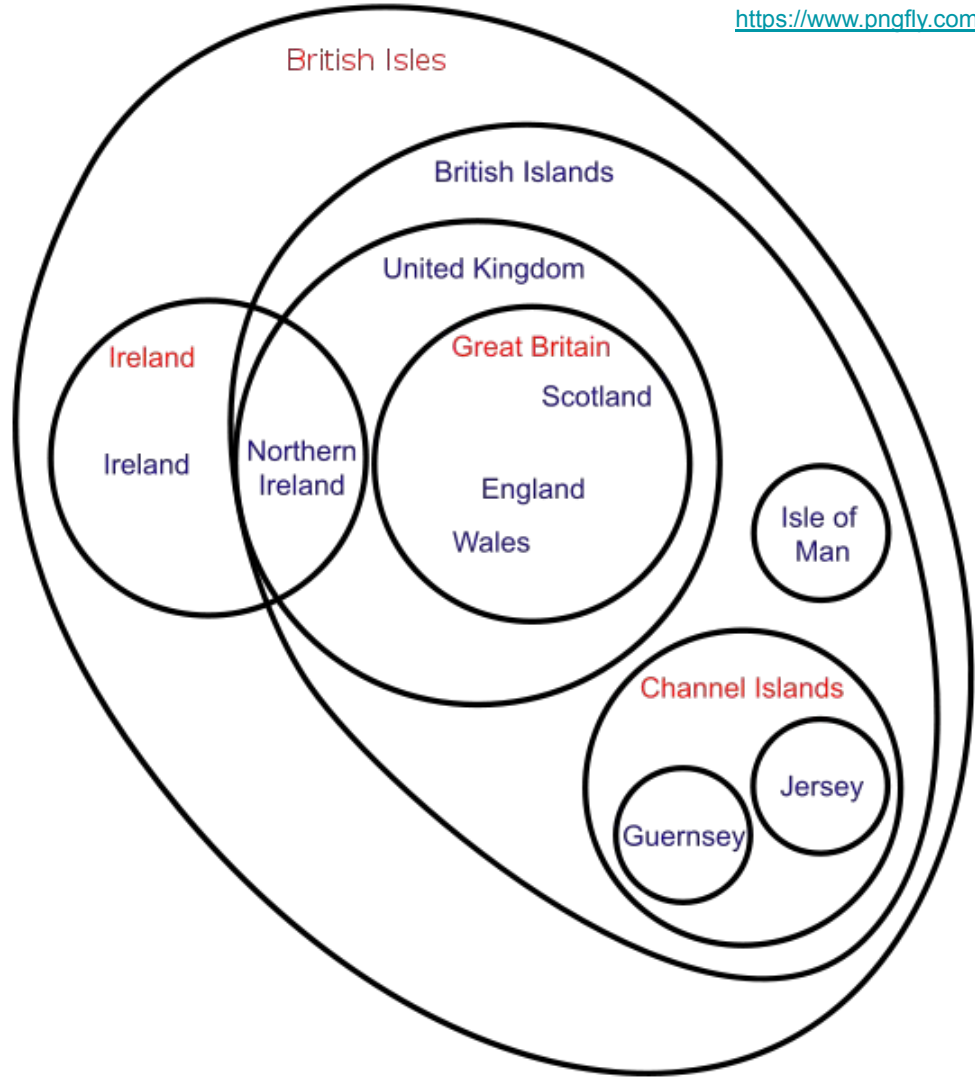
Misunderstood.

Maybe this Venn Diagram will explain this better :



Euler diagrams

Only show existing set relations.

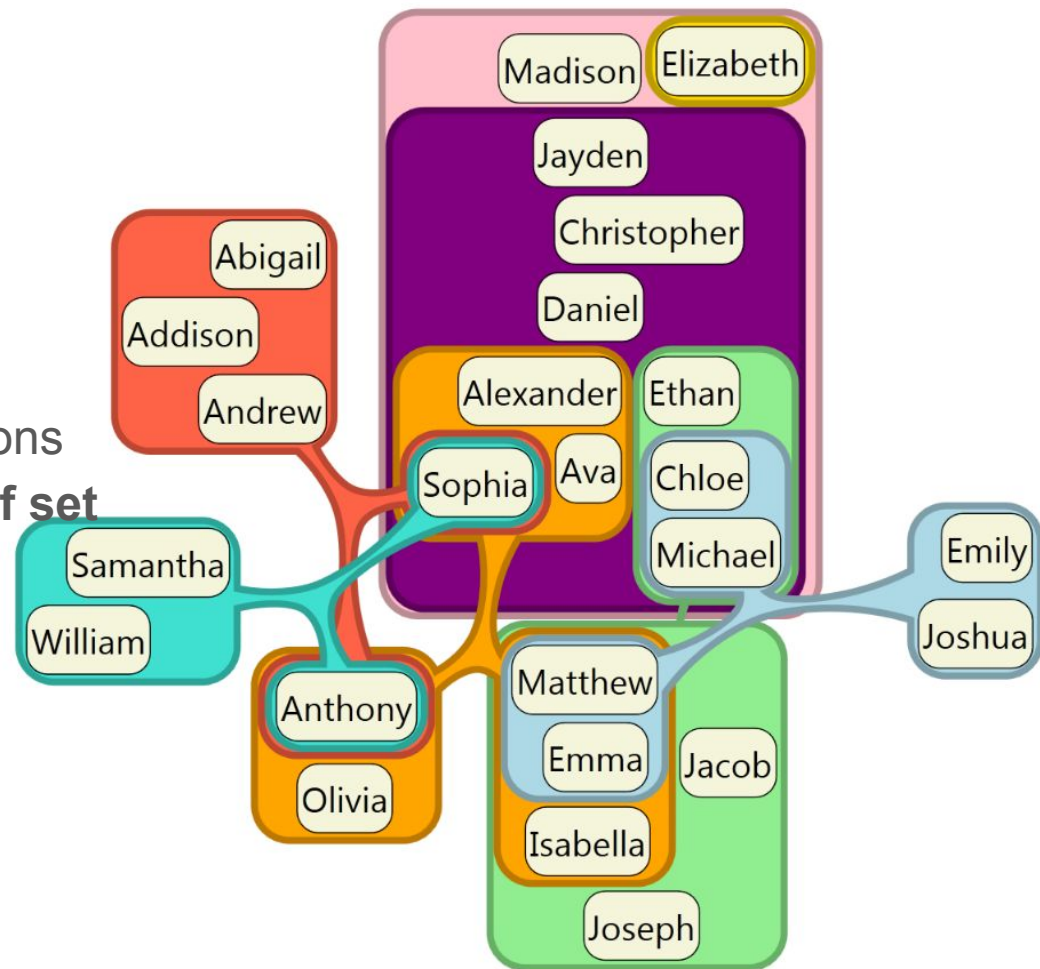


Euler diagram variants

Sets in rectangular shapes.

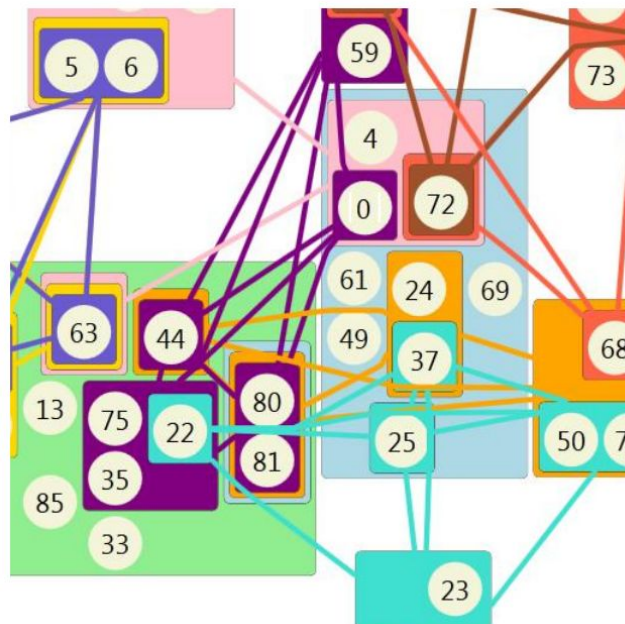
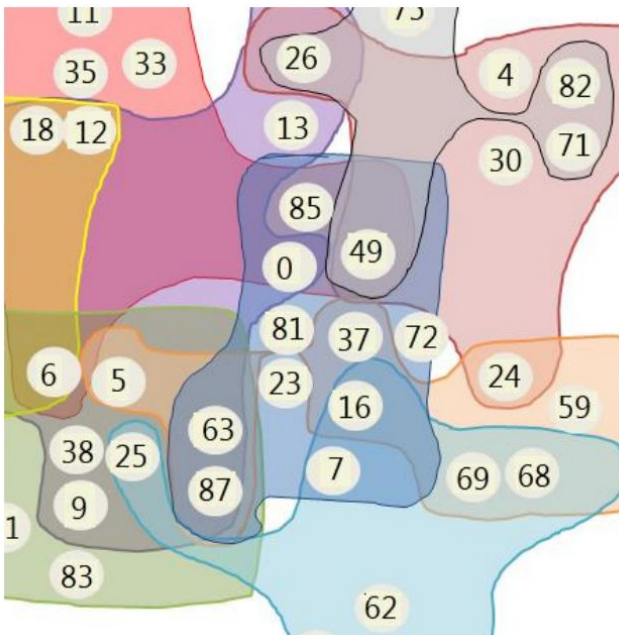
Split sets with intersections.

Improve readability of set intersections by **minimizing the intersections of set regions' boundaries**.



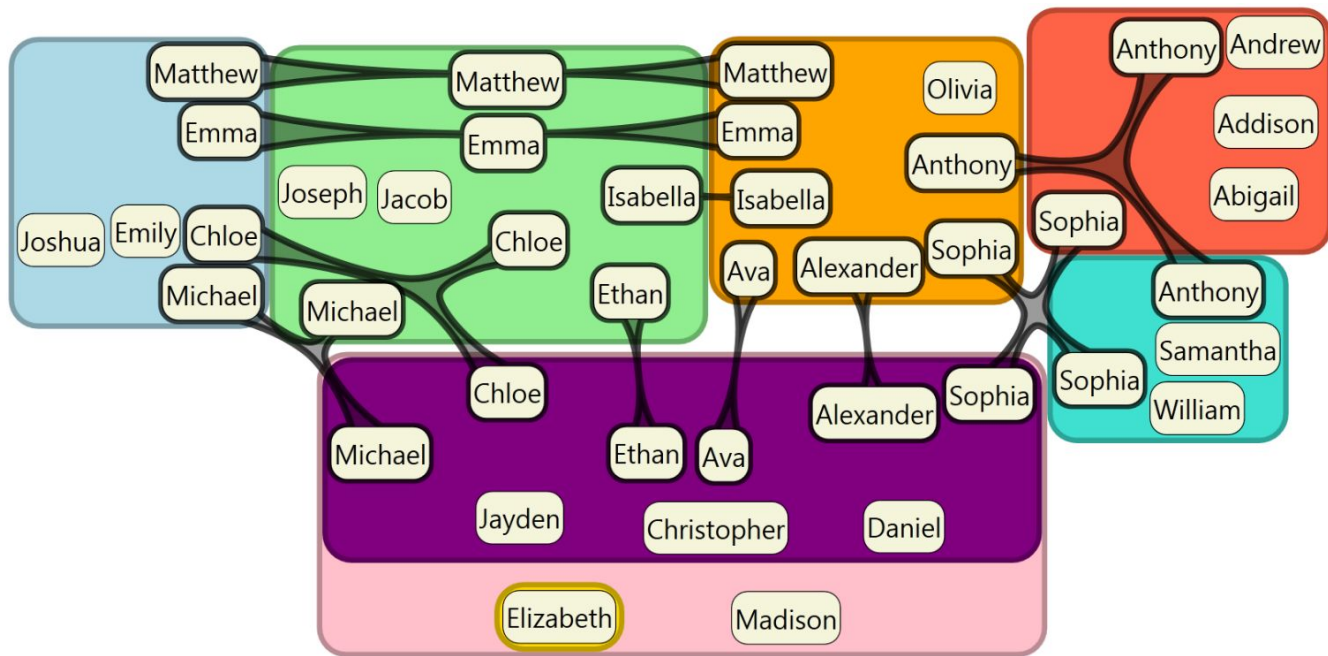
Euler diagram variants

Sets in rectangular shapes.



Euler diagram variants

Duplicate elements belonging to multiple sets.



Overlays

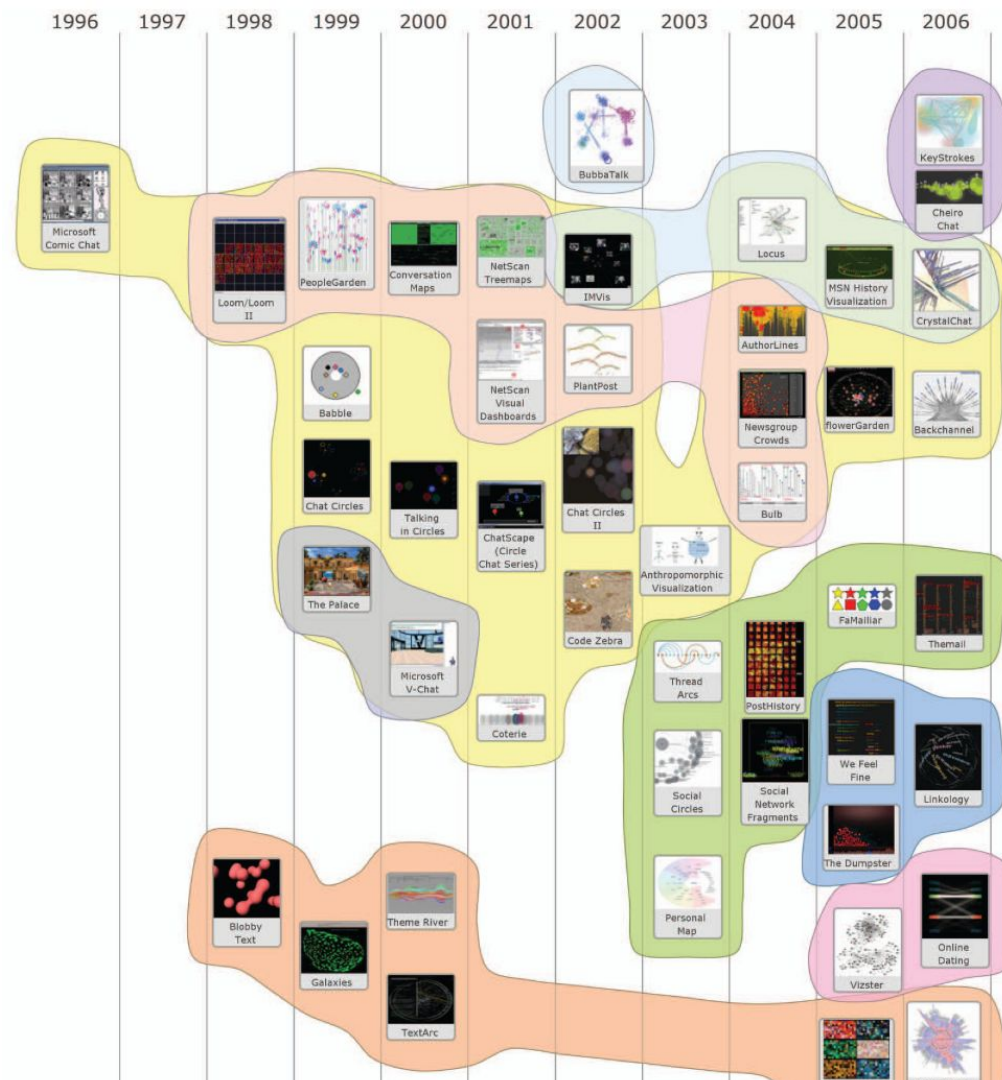
Augment set memberships over the elements in an existing visualization.



Overlays: Bubble Sets

Overlay a set membership relation on top of a primary data relation.

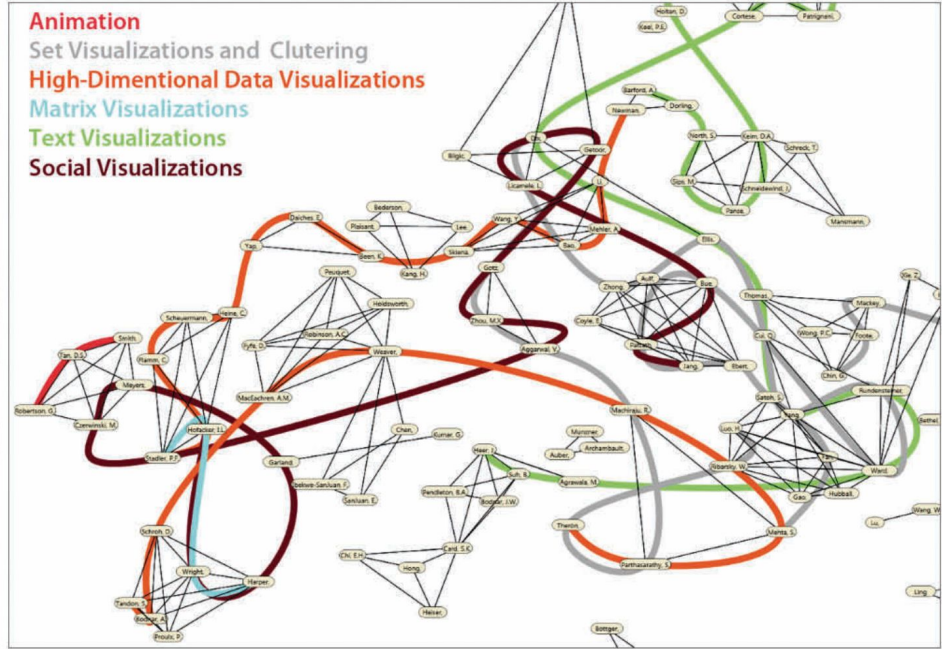
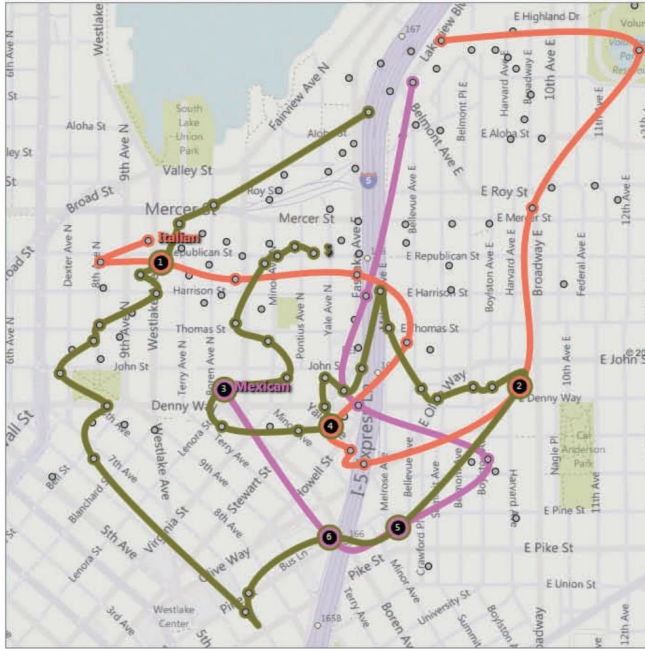
Use line contours.





Overlays: LineSets

One smooth line to connect elements of one set with minimal length. This in turn avoids self-crossings.

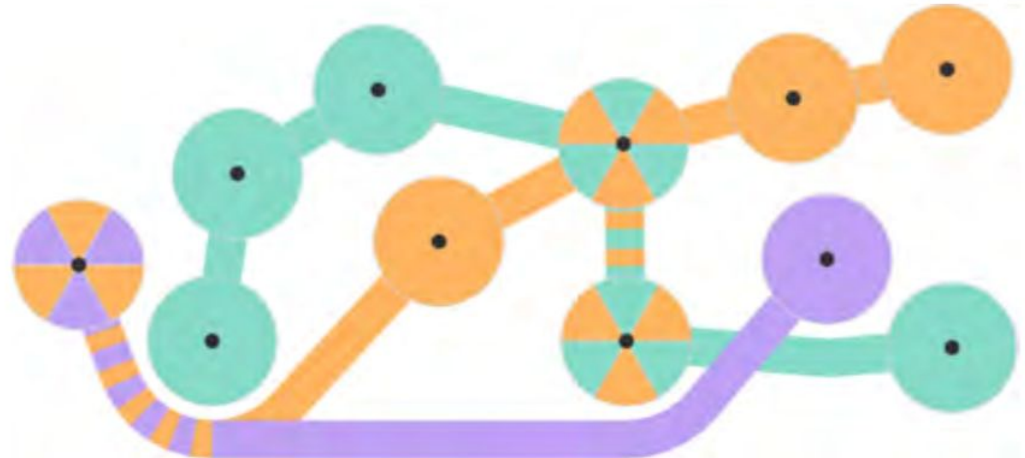
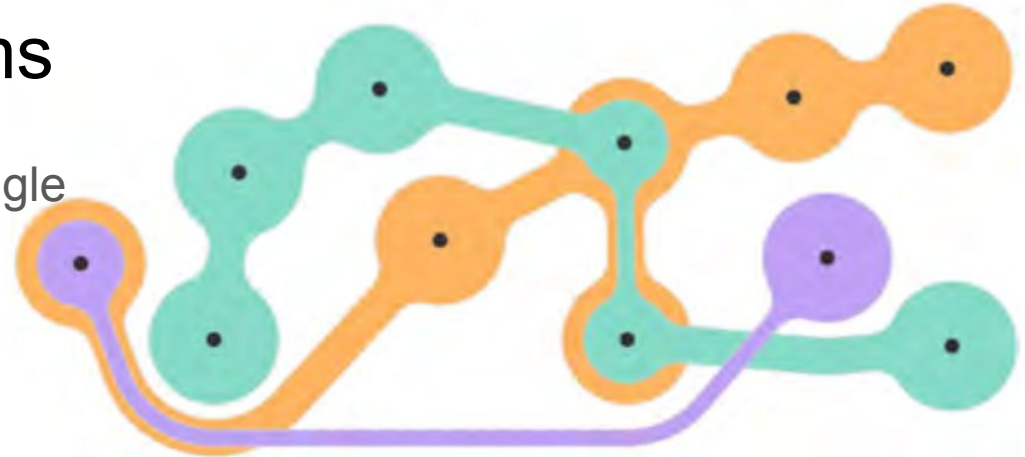


Overlays: Kelp Diagrams

A **graph structure**, rather than a single line, to connect elements of a set.

Nested style

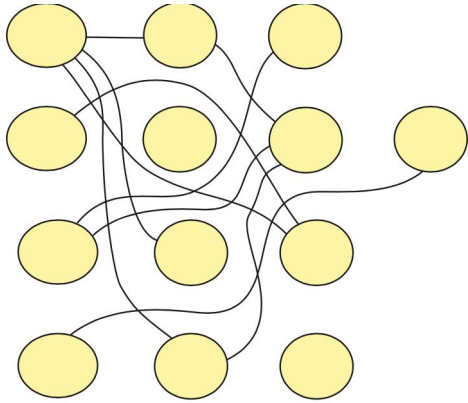
Striped style



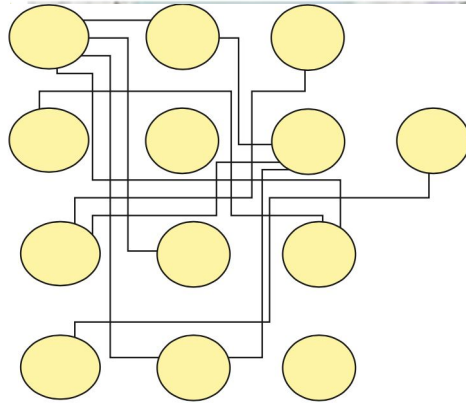
Overlays: Comparison

LineSets has more bends, and line crossings.

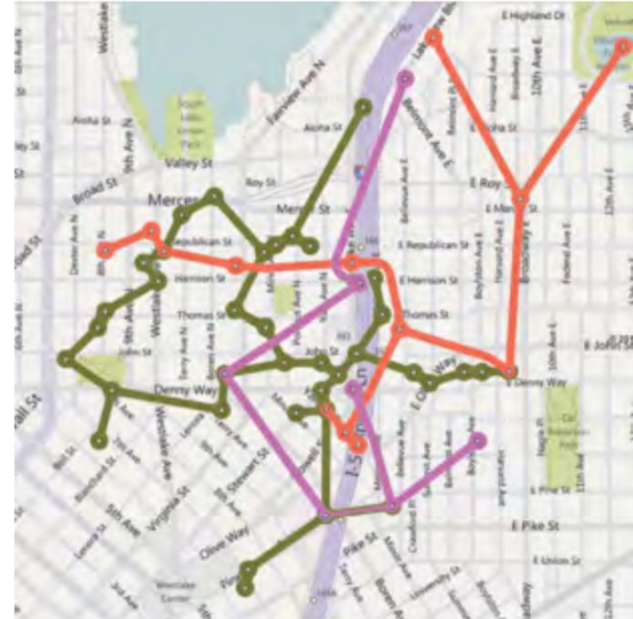
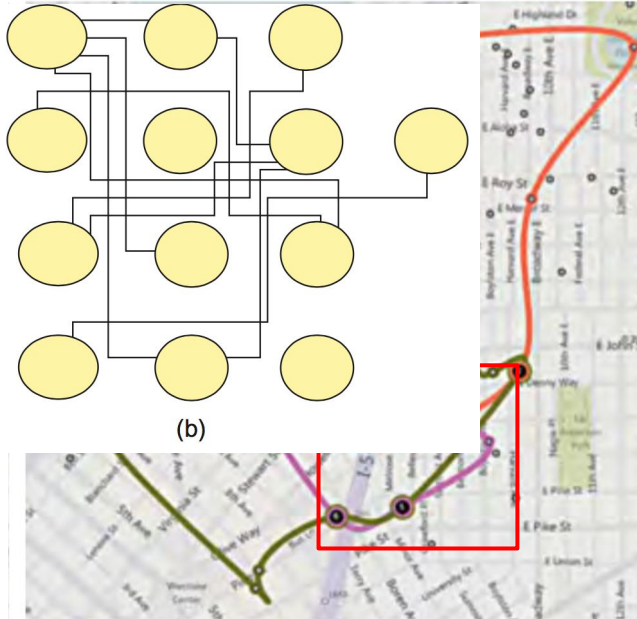
Kelp Diagrams are 1) computationally too slow for interaction;
2) not easy to visually follow line connections (Gestalt Laws of continuity).



(a)



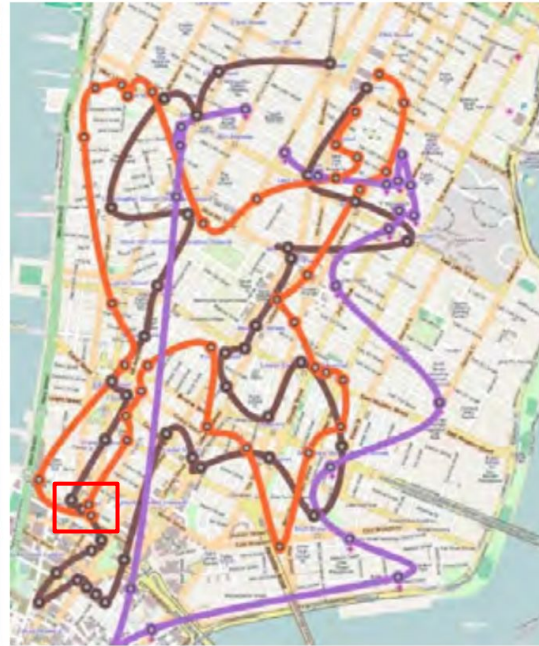
(b)



Overlays: Comparison

Bubble Sets: color blending introduces new colors. 4-20 sets.

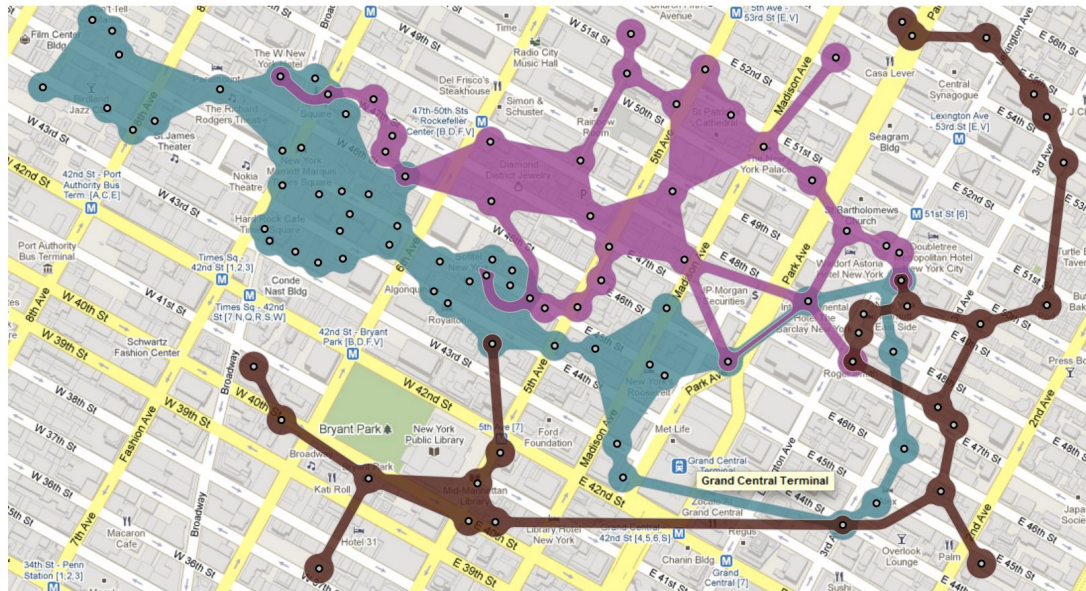
LineSets: not try to route around the elements not belonging to a set. Tens of sets.



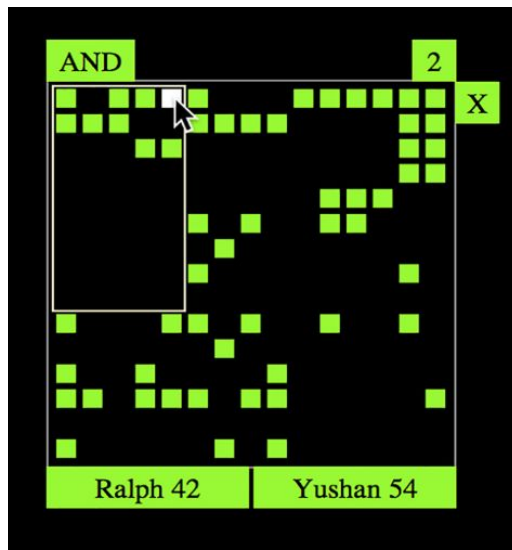
Overlays: Comparison

LineSets and Kelp Diagrams: complex contours for clusters of elements.

KelpFusion: lines + regions



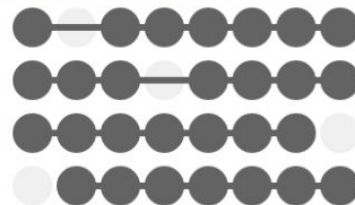
Matrix-based techniques



416-Men...underpants
416-Men...suits knit.
416-Men...singlets
416-Men...crocheted
416-Men...suits
416-Men...shirts
416-Men...ov. others
416-Men...overcoats

+ Query

▼ Degree 7 (7 set intersect.)



▼ Degree 8 (8 set intersect.)



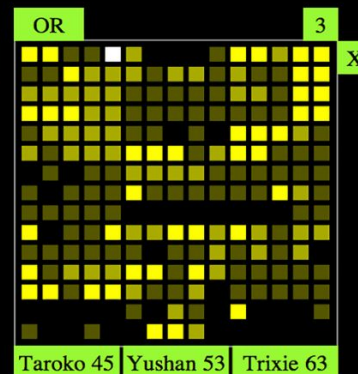
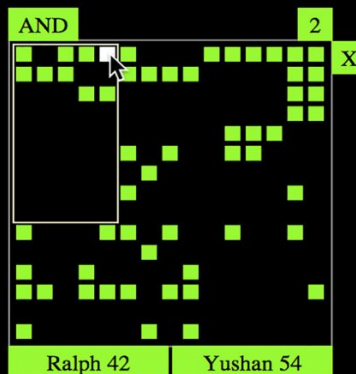
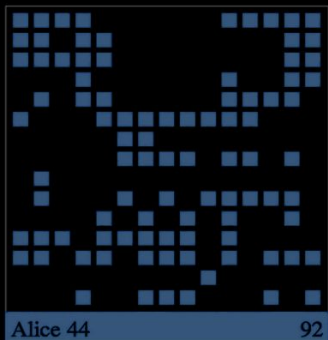
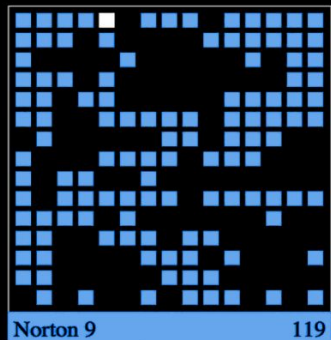
Matrix-based: OnSet

Guanidoacetic acid

Amino Acids

Samples

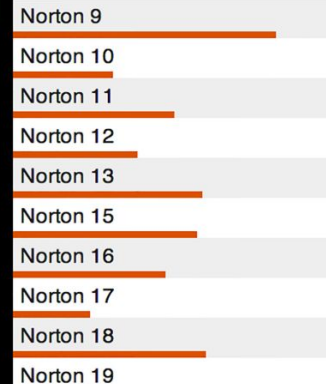
Compounds



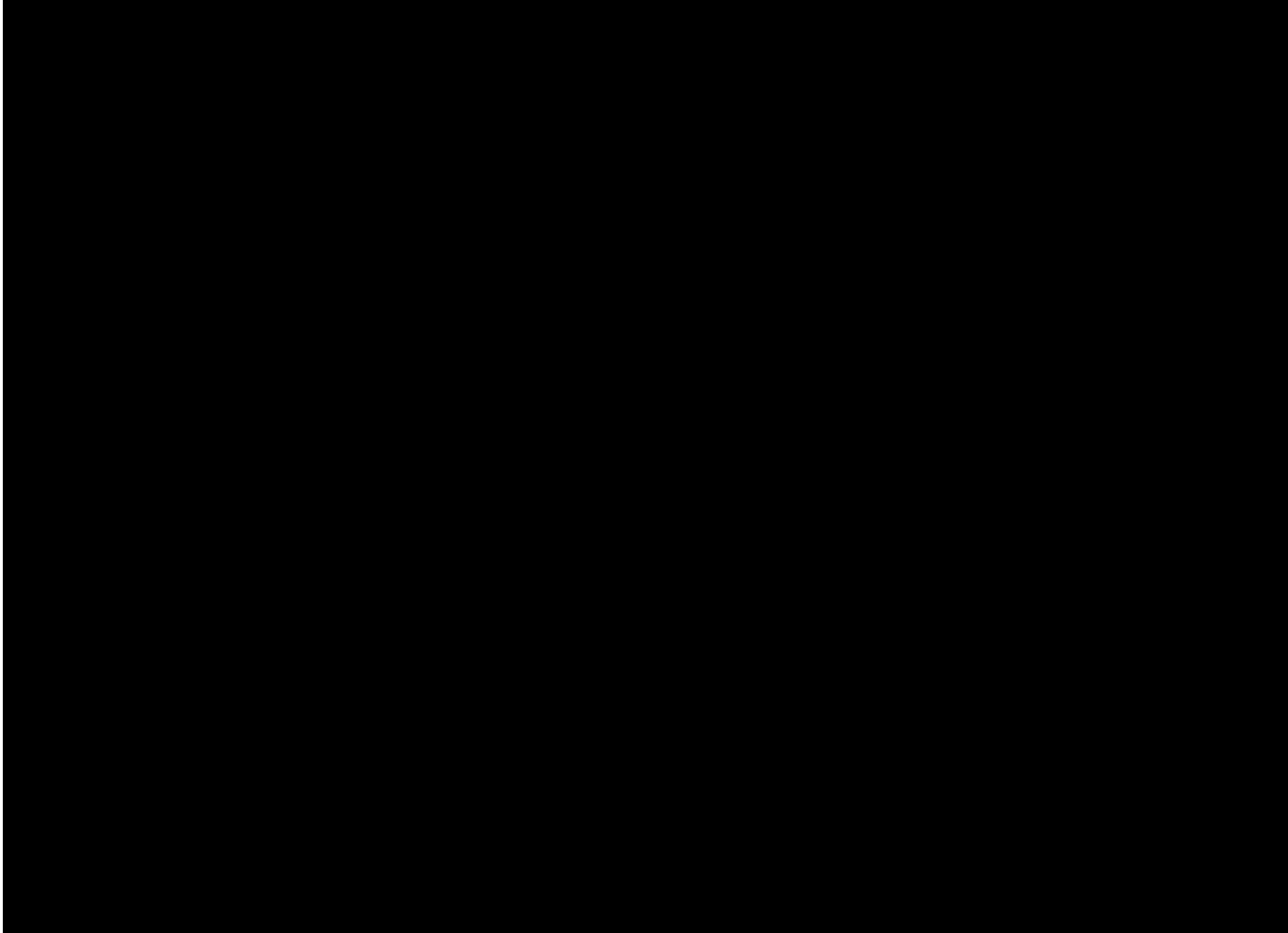
Search

Def.

A→Z



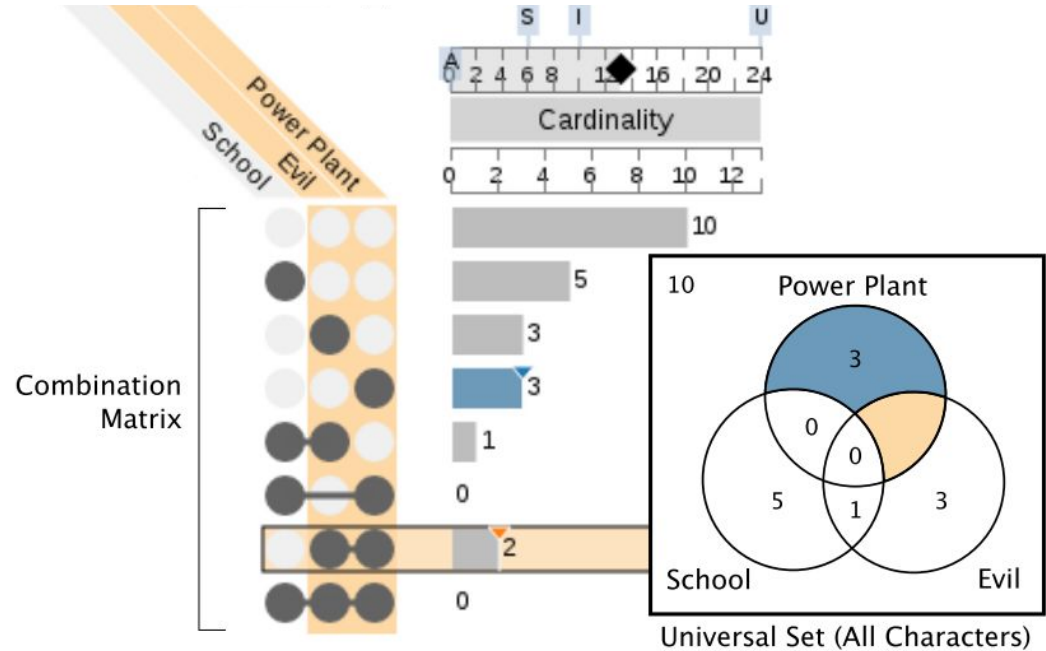
Sadana, R., Major, T., Dove, A. and Stasko, J., 2014. Onset: A visualization technique for large-scale binary set data. *IEEE transactions on visualization and computer graphics*, 20(12), pp.1993-2002.

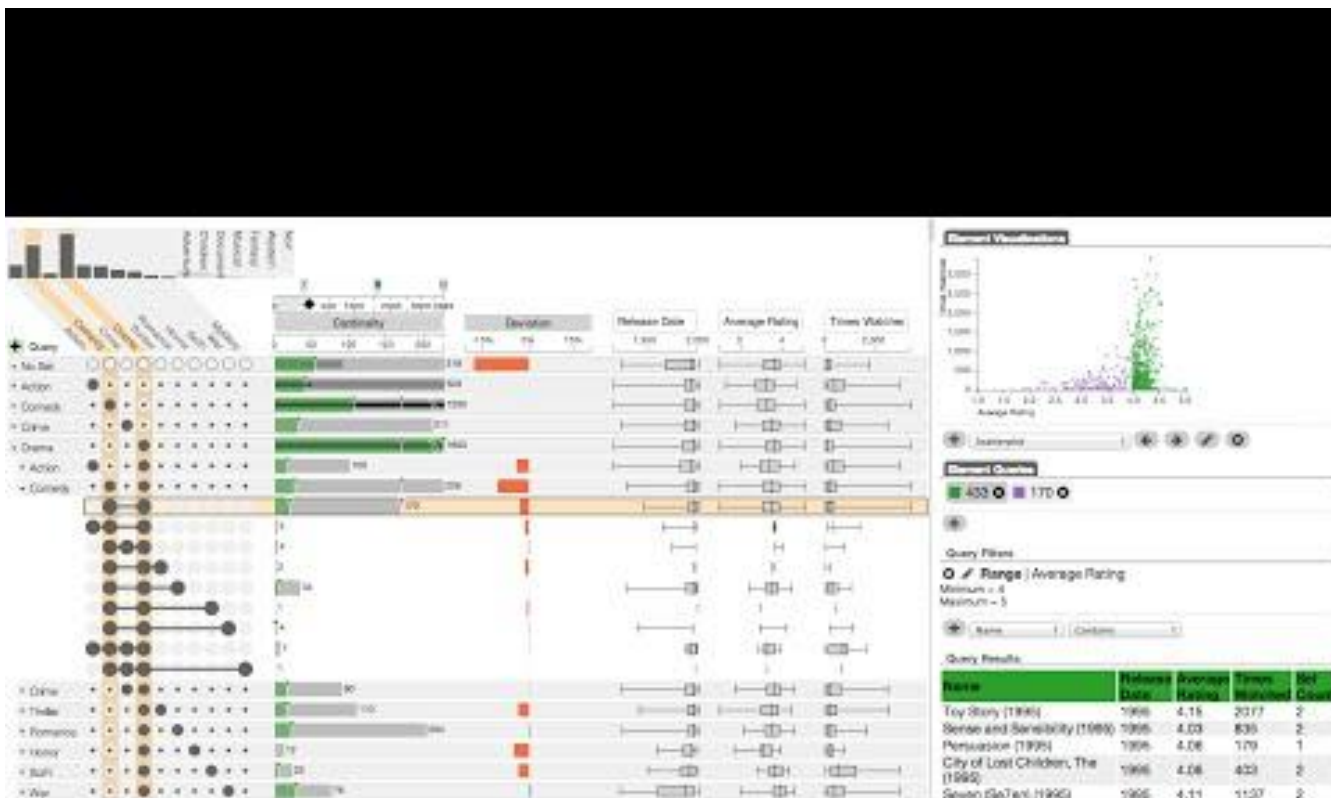


Matrix-based: UpSet

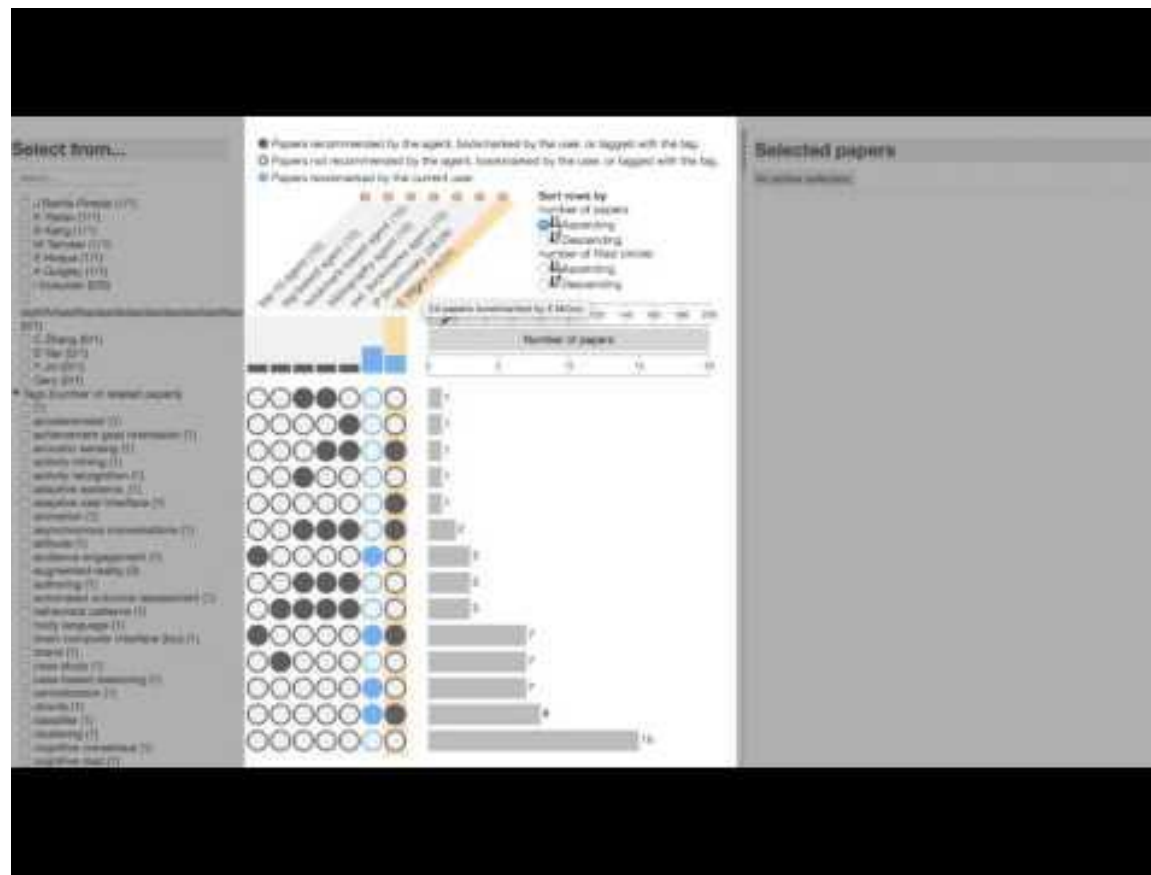
Row: set relation

Column: set



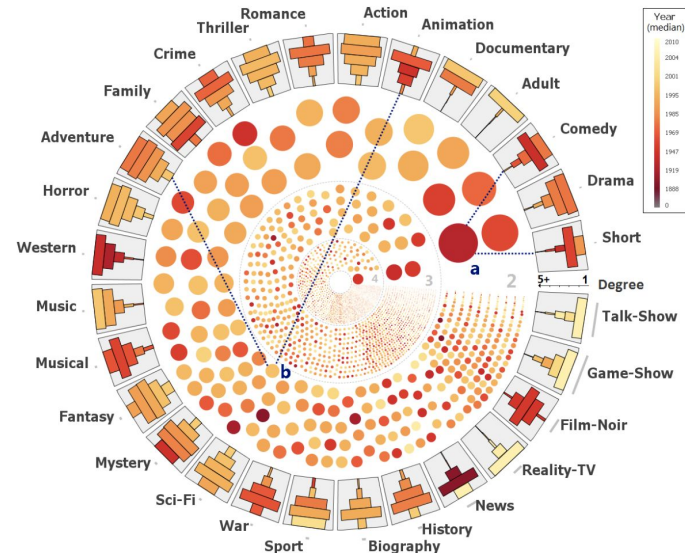
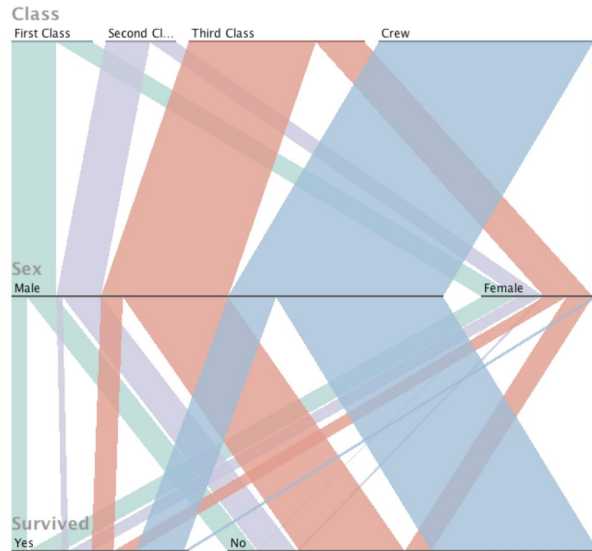


UpSet application: exploring publications from intersecting prospects



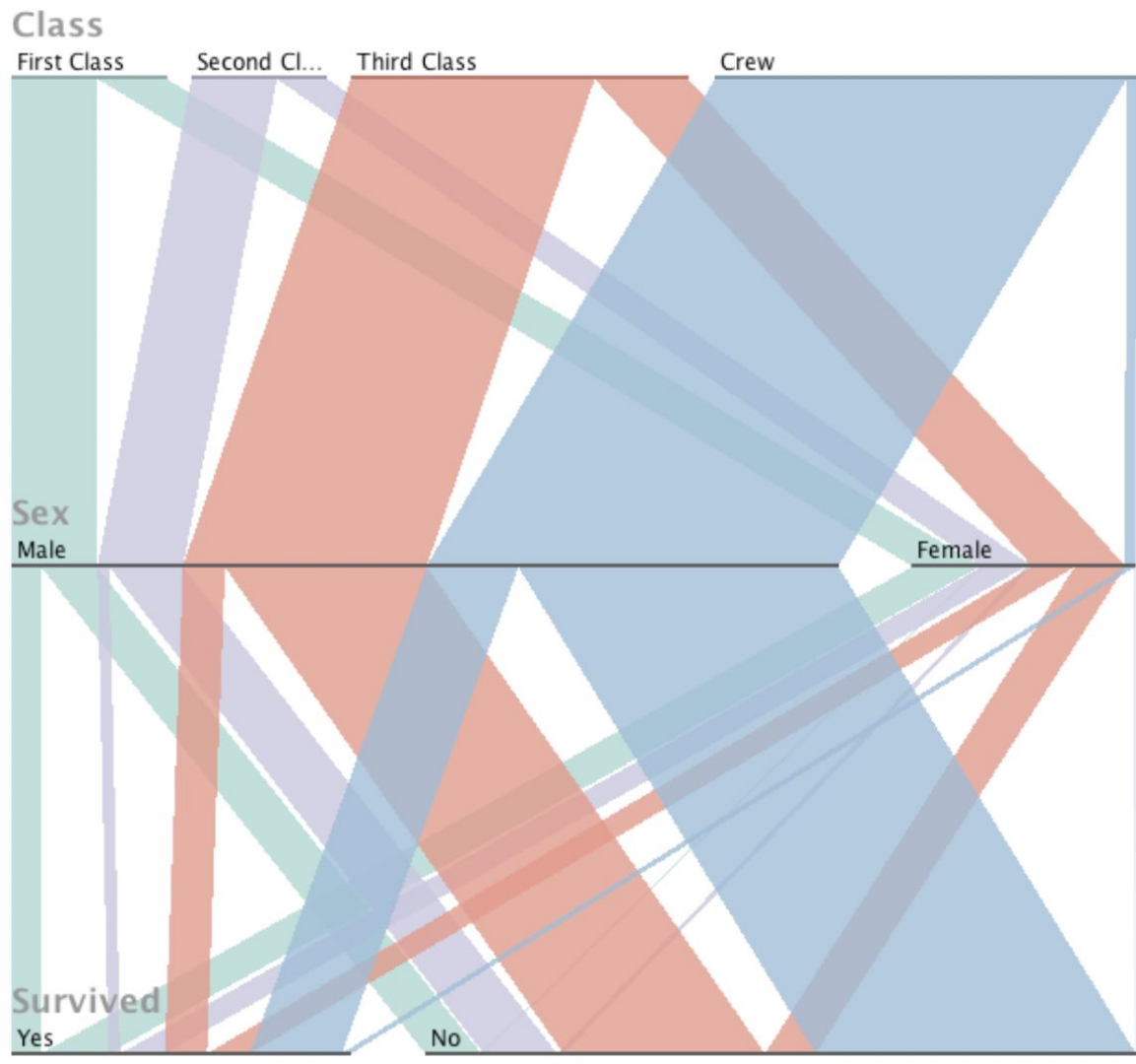
Aggregation-based techniques

Too many elements: Aggregate data elements into a single visual element that encodes the frequency.

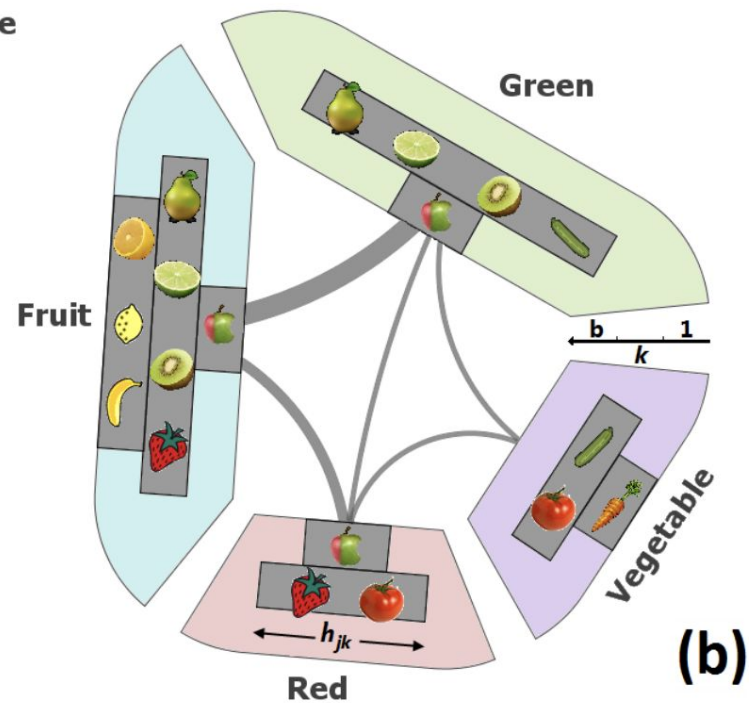
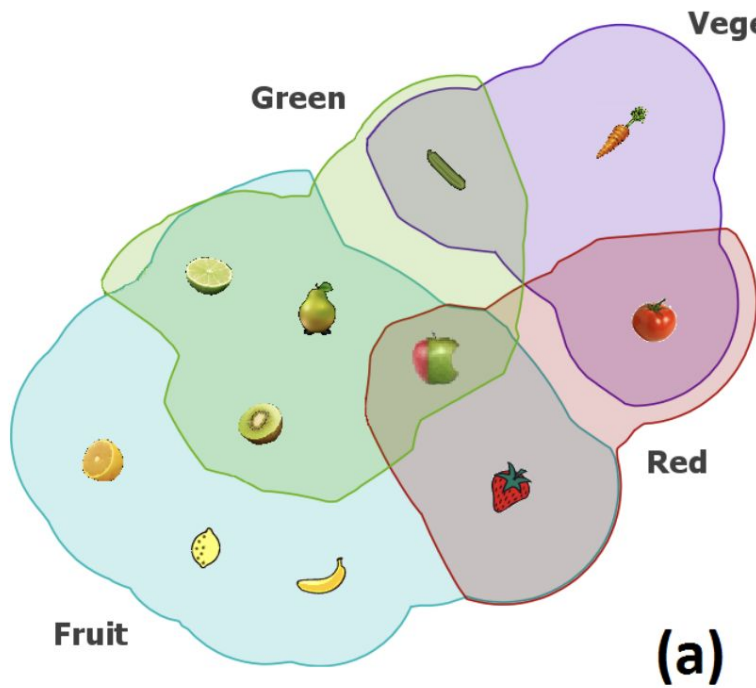


Aggregation-based: Parallel sets

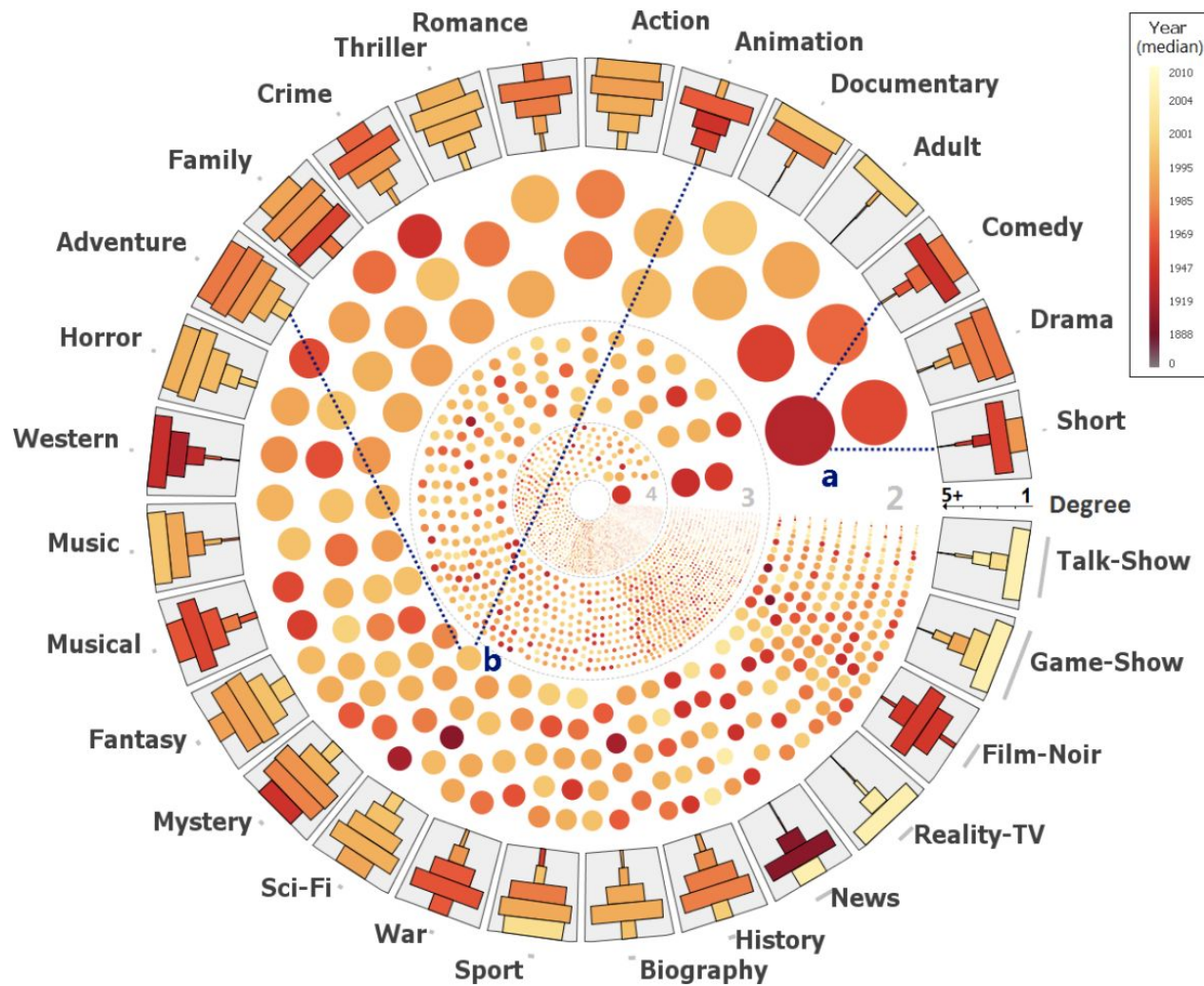
Show data frequencies rather than individual data points.

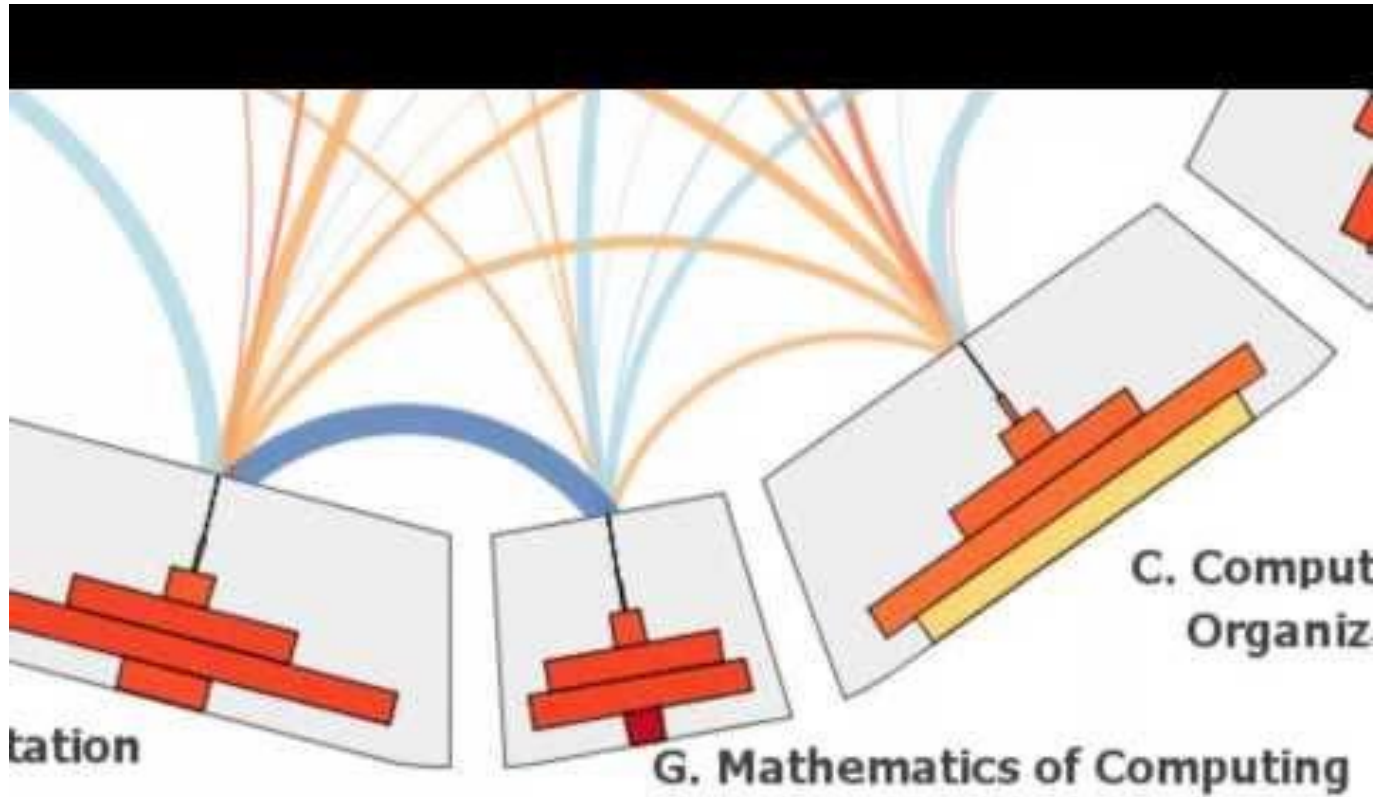


Aggregation-based: Radial sets



Radial sets

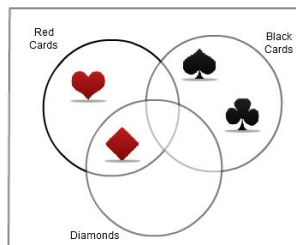




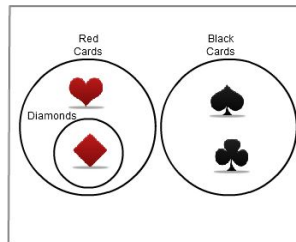
Recap

Venn and Euler diagrams

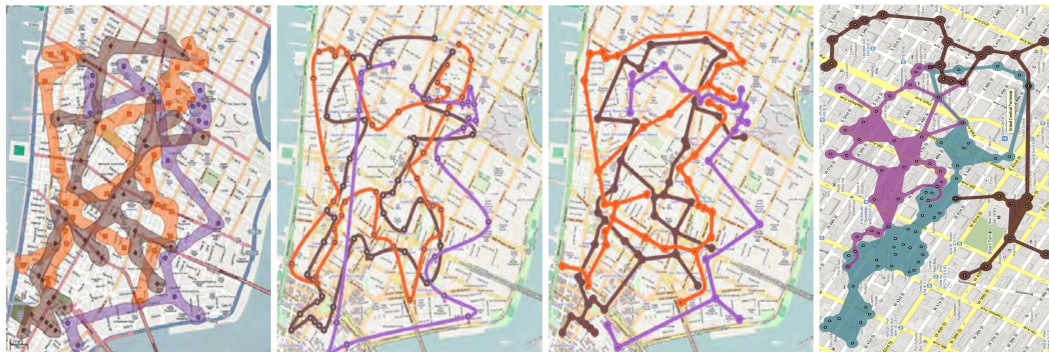
V
E
N
N



E
U
L
E
R

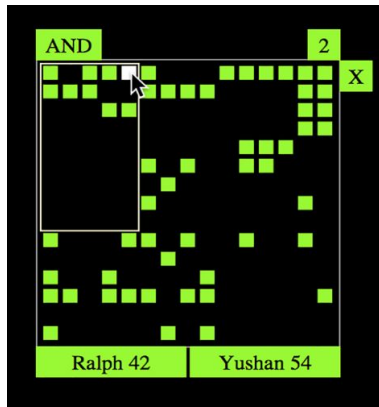


Overlays



Recap

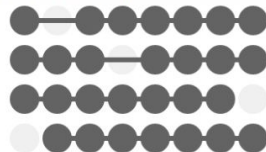
Matrix-based



416-Men...underpants
416-Men...suits knit.
416-Men...suits
416-Men...singlets
416-Men...crocheted
416-Men...shirts
416-Men...ov. others
416-Men...overcoats

+ Query

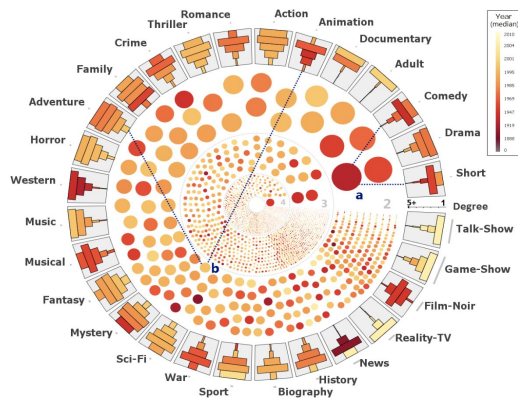
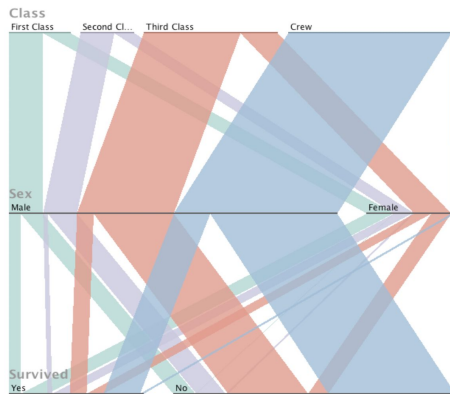
▼ Degree 7 (7 set intersect.)



▼ Degree 8 (8 set intersect.)



Aggregation-based

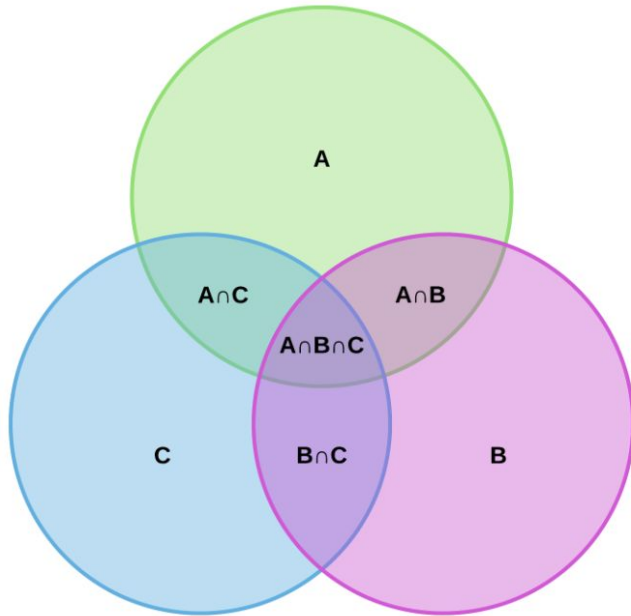


Which technique to use?

Intuitiveness

VS.

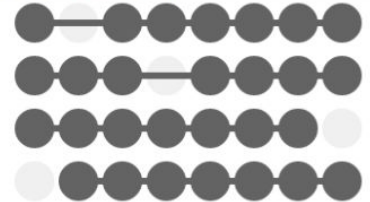
Scalability



+ Query

416-Men...overcoats
416-Men...ov. others
416-Men...shirts
416-Men...crocheted
416-Men...singlets
416-Men...suits
416-Men...suits knit.
416-Men...underpants

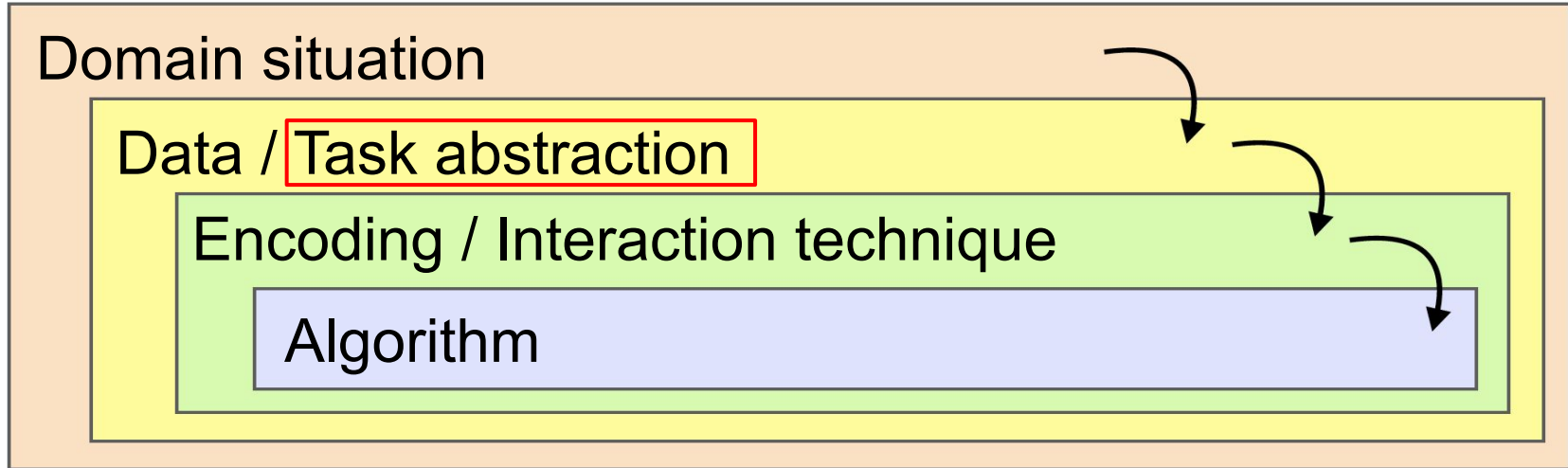
▼ Degree 7 (7 set intersect.)



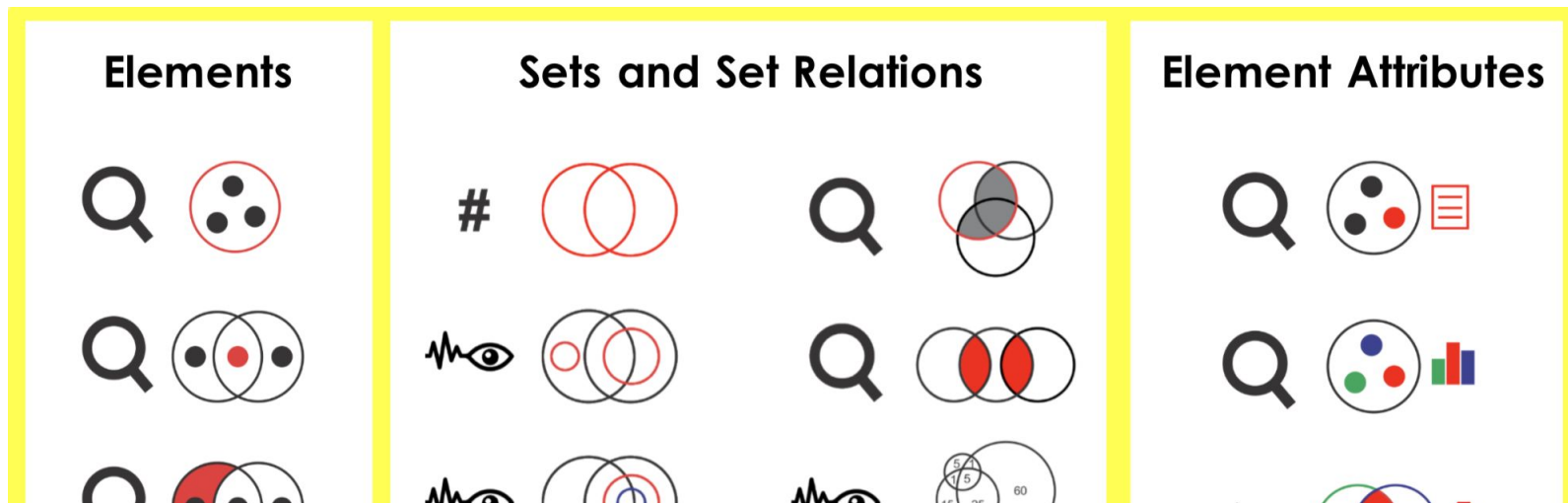
▼ Degree 8 (8 set intersect.)



Which technique to use?



Task abstraction



Alsallakh, B., Micallef, L., Aigner, W., Hauser, H., Miksch, S. and Rodgers, P., 2014, June. Visualizing sets and set-typed data: State-of-the-art and future challenges. In *Eurographics conference on Visualization (EuroVis)–State of The Art Reports* (pp. 1-21). The Eurographics Association.

Alsallakh, B., Micallef, L., Aigner, W., Hauser, H., Miksch, S. and Rodgers, P., 2016, February. The State-of-the-Art of Set Visualization. In *Computer Graphics Forum* (Vol. 35, No. 1, pp. 234-260).