

The 22nd International Conference on Asia-Pacific Digital Libraries (ICADL) 2020

November 30, 2020, On-line

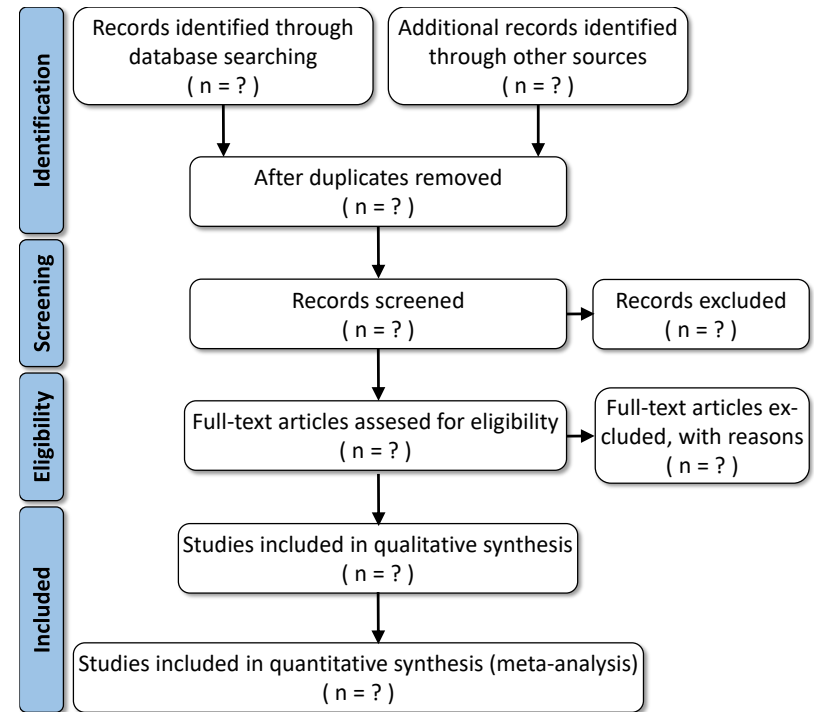
ReViz: A Tool for Automatically Generating Citation Graphs and Variants

Professor Dr. rer. nat. habil. Sven Groppe

<https://www.ifis.uni-luebeck.de/index.php?id=groppe>

Systematic Reviews

- **Aim:** Search for relevant scientific contributions of an area of interest in a systematic way
- Result of search is often illustrated in a flow chart
- Analysis of relevant contributions published in survey paper
- **Manual** Systematic Reviews cause a lot of efforts & are prone to errors
- → Tool support for systematic reviews necessary
- → **Our focus:** Automatic visualizations for systematic reviews



Visualizations in Survey Papers

SLR	Flow chart: selection of publi- cations	Table Study pro- perties	Other Tables	Distri- bution of publ.- years	Other distri- butions	Fo- rest Plot	Fun- nel chart	Other
[13]	✓		✓			✓	✓	
[32]	✓		✓			✓		
[25]	✓					✓	✓	
[18]	✓	✓				✓		Risk of bias diagram
[33]	✓			✓	✓			classifications of relevant publications
[3]			✓	✓				
[4]	✓	✓	✓	✓				
[26]	✓		✓	✓	✓			
[10]	✓		✓					
[8]	✓	✓	✓					

- Most survey papers contain numerous visualizations for analysis

Tool Support for Systematic Reviews - Example Parsifal

Parsifal Blog About Help LinaH ⚙️ 🔄

LinaH / Systematic literature review [Review settings](#)

Review Planning **Conducting** Reporting

1. Search 2. Import Studies 3. Study Selection 4. Quality Assessment 5. Data Extraction 6. Data Analysis

Study Selection

All Sources Science@Direct Scopus

Action: 0 of 15 selected

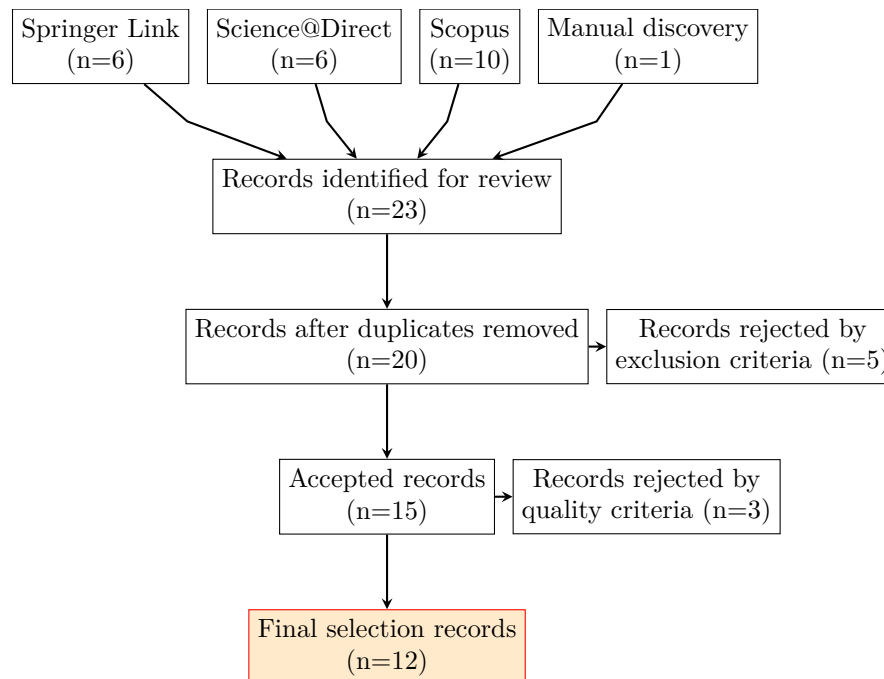
Show: ☒ All ☐ Accepted ☐ Rejected ☐ Unclassified ☐ Duplicated

☐	Bibtex Key	Title	Author	Journal	Year	Added by	Added at	Status
☐	kitchenham2004procedures	Procedures for performing systematic reviews	Kitchenham, Barbara	Keele, UK, Keele University	2004	LinaH	06 Dec 2019 10:47:56	Rejected
☐	prisma	Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement	Moher, David and Liberati, Alessandro and Tetzlaff, Jennifer and Altman, Douglas G. and the PRISMA Group	Annals of Internal Medicine	2009	LinaH	06 Dec 2019 10:47:56	Unclassified
☐	ALZUBIDY201772	Vision for SLR tooling infrastructure: Prioritizing value-added requirements	None	Information and Software Technology	2017	LinaH	06 Dec 2019 10:47:56	Accepted

- **No Support of Automatic Generation & Export of Visualizations**

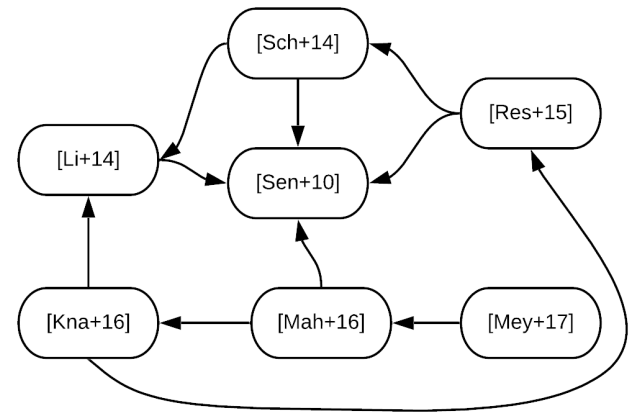
ReViz: Features: Flow chart for the Selection of Publications

- **Output: tikz code** for easy integration into latex papers
 - supports easy manual modifications



Citation Graphs

- Nodes represent **publications**, edges represent **citation relationships**
- **Analysis**
 - Spread of ideas & knowledge
 - Influences of contributions
 - Impact of publications
 - Detection of **communities**, author circles, self citations
- **Drawbacks** of simple form of citation graphs
 - Could be more obvious:
 - Publication history
 - Number of (in)direct citations
 - Publications with common authors
 - Complexity of graphs increases with number of publications

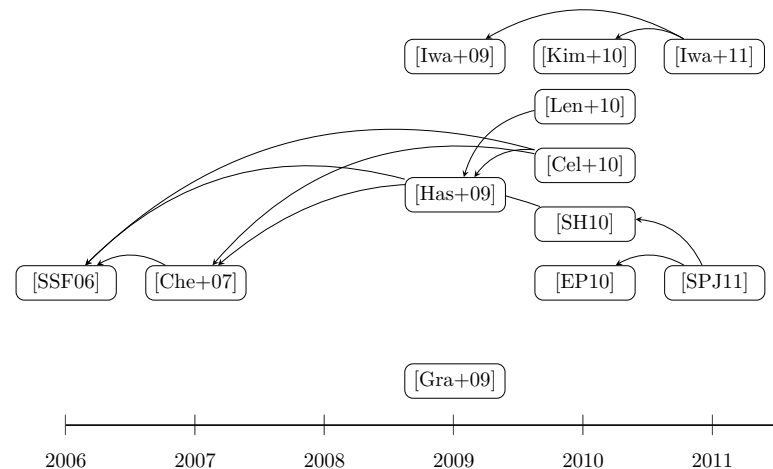


ReViz: Features 1/8: Timeline in Cit. Graphs

- **Output: tikz code** for easy integration into **latex papers**
 - supports easy manual modifications
- **Timeline**
 - **Publication history**
 - **Edges (typically) from right to left**
 - **Spread of ideas and knowledge from left to right:**
Which publications can be reached from a publication?

ReViz: Features 2/8: Isolated Subgraphs in Cit. Graphs

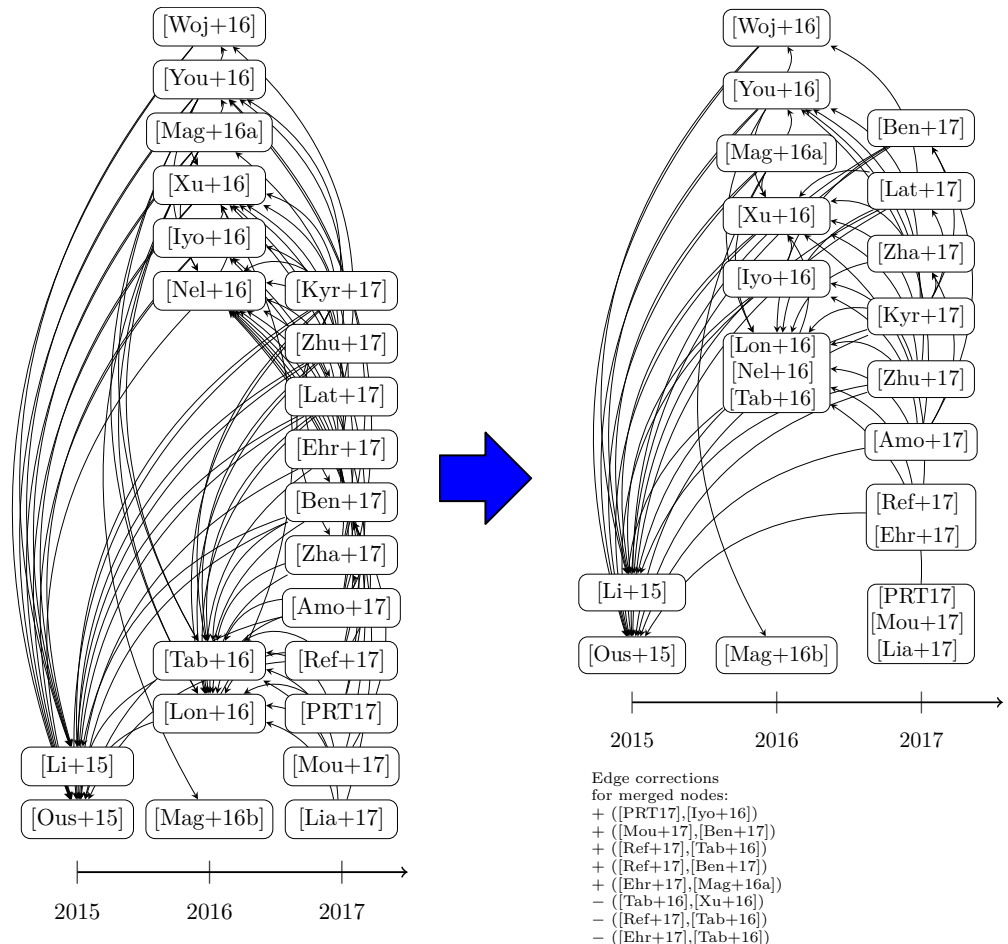
- Isolated subgraphs are displayed **separately** on different layers
 - Publications dealing with different topics, or
 - Publications dealing with the same/similar topics, but **from different communities**



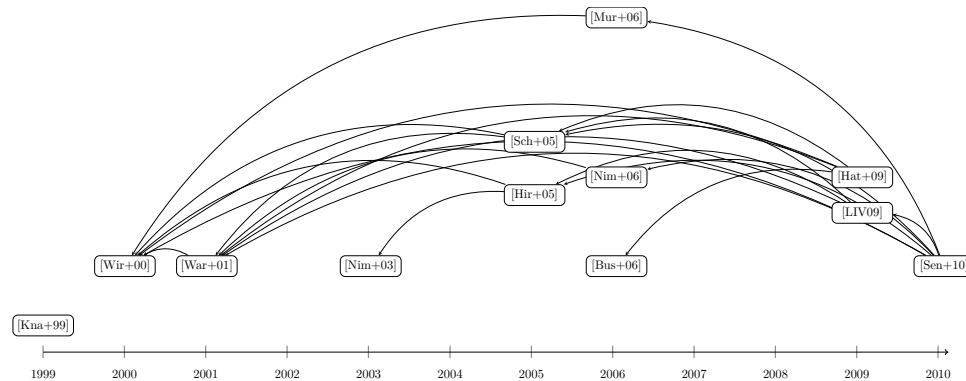
- Node layout according to approach of Sugiyama¹
 - Proven method in order to avoid crossing edges as much as possible

ReViz: Features 3/8: Node Summaries

- Simplification of citation graphs
- Merging of nodes with edges from/to common nodes
- Edge corrections when merging nodes, where only subset of edges are common



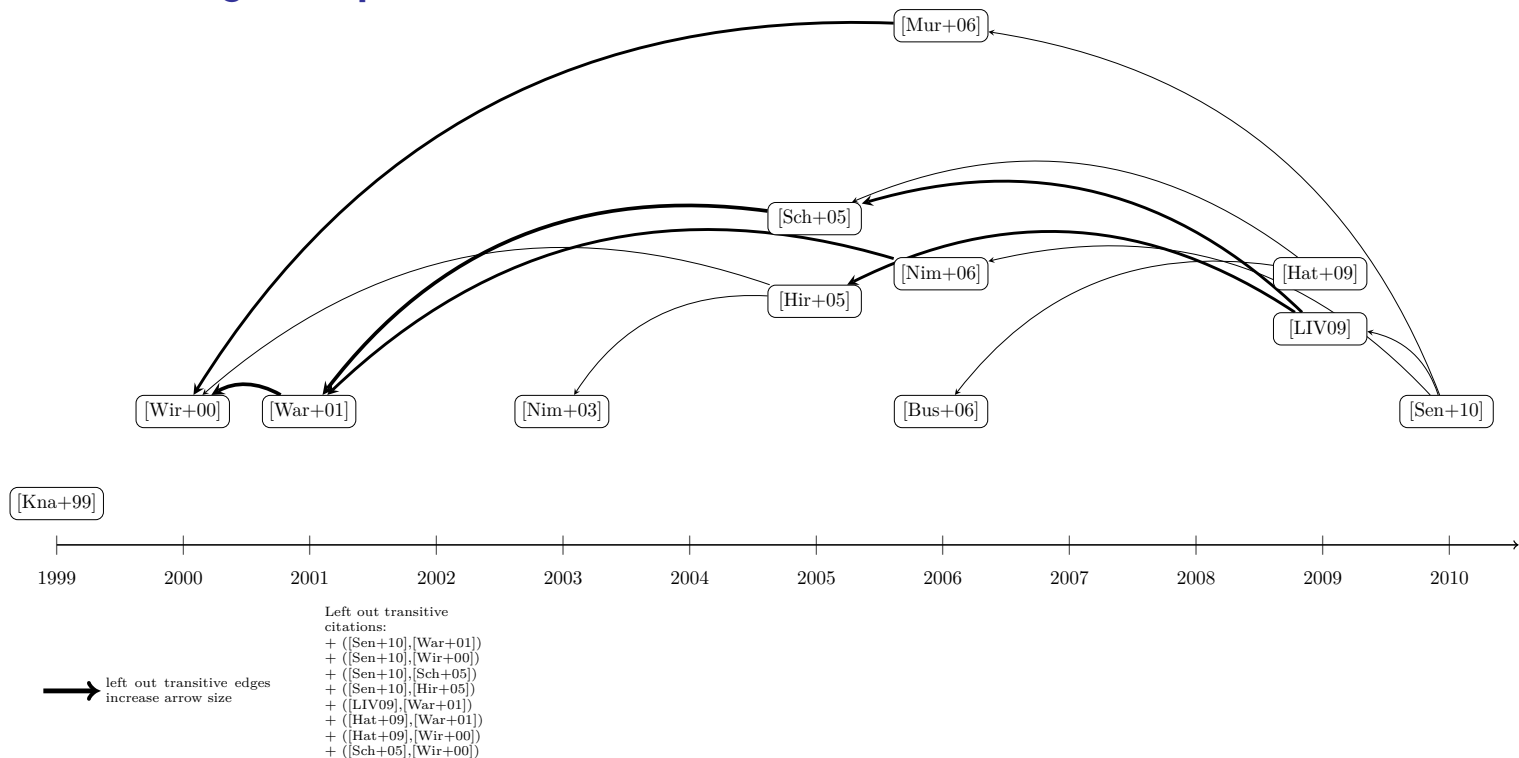
ReViz: Features 4/8: Transitive Edges 1/2



- Reduce graph complexity by removing transitive edges
 - Drawback: Remaining edges $\neq$ number of direct citations

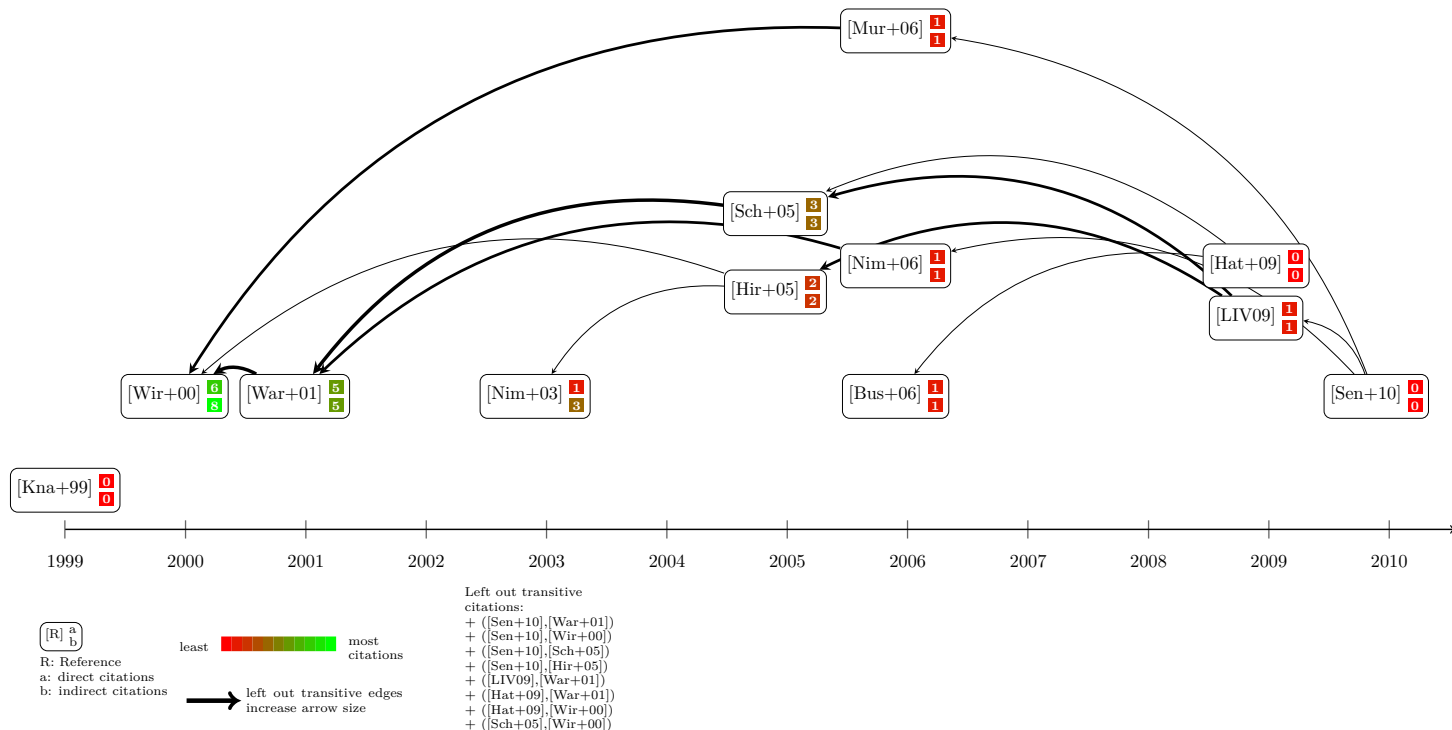
ReViz: Features 5/8: Transitive Edges 2/2

- **More** removed transitive $\Rightarrow$ **larger** width of ingoing edges
 - **Increasing width** of ingoing edges for
 - **more citations**
 - **higher impact**



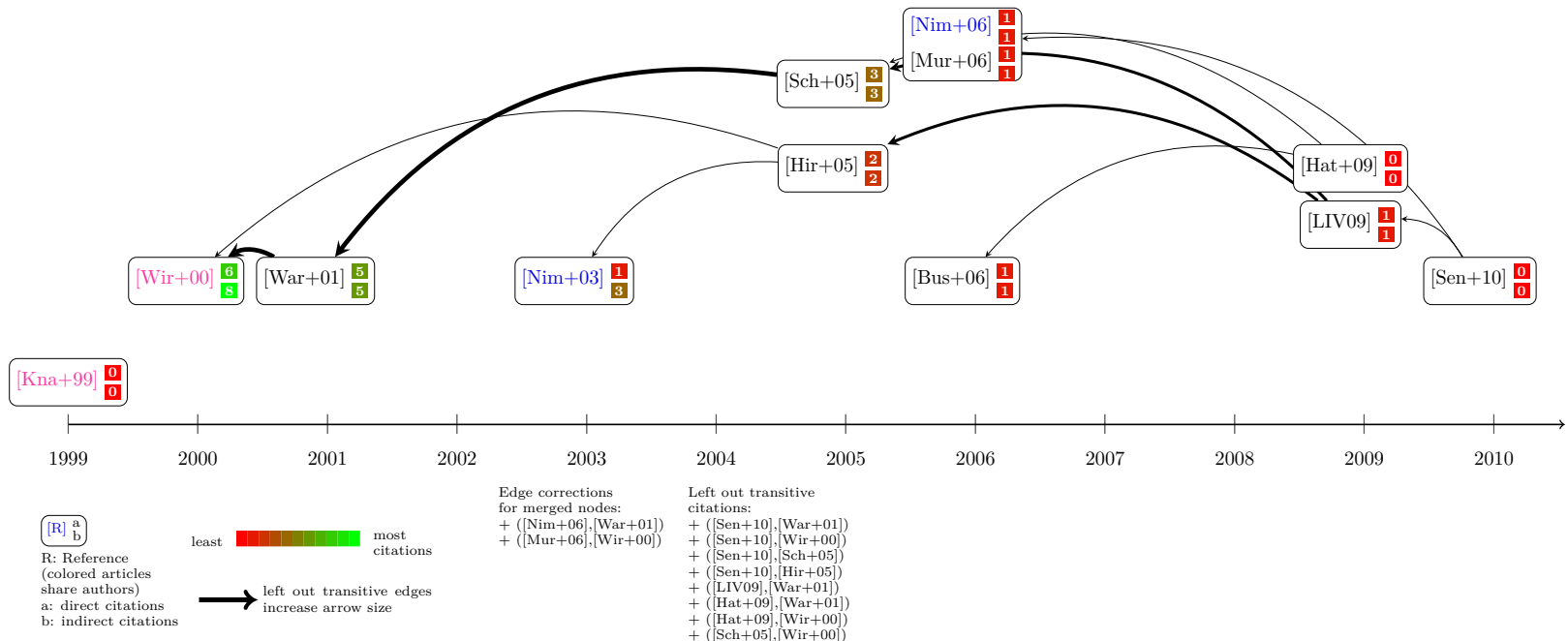
ReViz: Features 6/8: # (In-)Direct Citations

- Displaying direct & indirect citations within node labels
 - Numerical comparison of number of (in-) direct citations ($\rightarrow$ impact)
 - Use of color scale for fast recognition of high impact publications



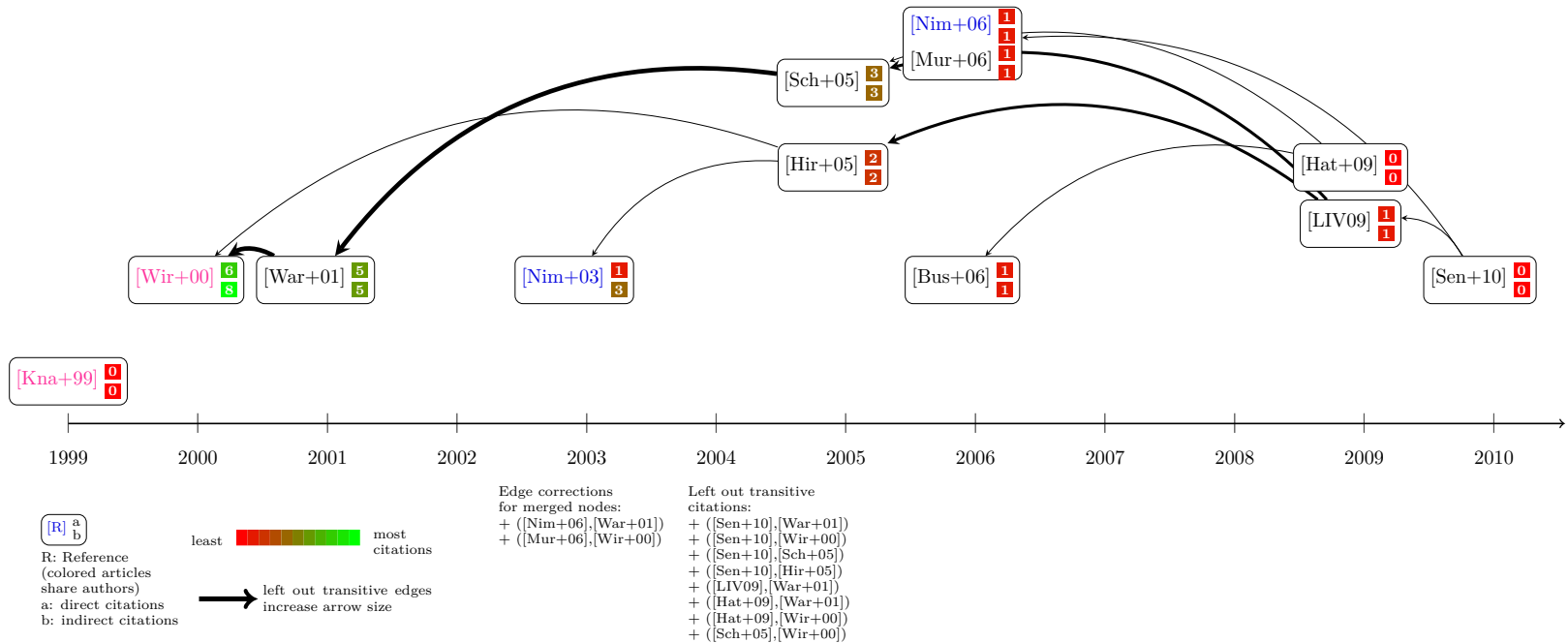
ReViz: Features 7/8: Common Authors

- Coloring publications with common authors with the same color
 - research groups become visible
 - Hint of follow-up papers & improved approaches of same researchers

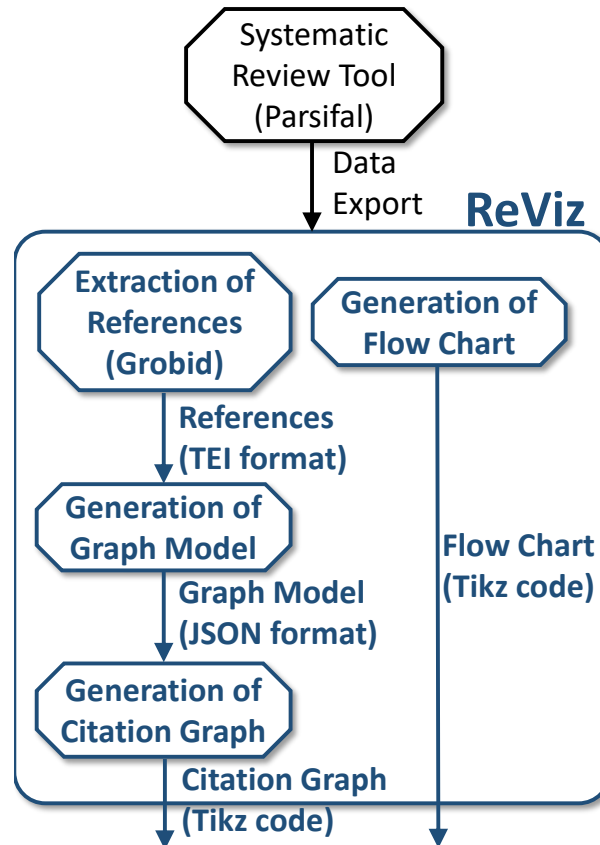


ReViz: Features 8/8: Features Optional

- All ReViz features can be switched on/off separately
- Example of all features on:



Architecture of our tool



- Parsifal is **optional**: Our tool works without it, too!
- Publicly **available**: <https://github.com/l-hartung/reviz>

Comparison with other tools

Visualiza- tion Tool	Visualiza- tions*	Citation graph from set of paper documents	Support of systematic reviews	Comments
CitNetEx- plorer [31]	C	-	-	Clustering of very large citation networks.
VOSViewer [30]	B	-	-	Visualized bibliometric networks constructed based on citation, bibliographic coupling, co-citation, or co-authorship relations may include journals, researchers, or individual publications.
Sci²Tool [5]	B	-	-	Temporal, geospatial, topical, and network analysis and visualization of scholarly datasets at the micro (individual), meso (local), and macro (global) levels.
CiteSpace [7]	B	-	-	Structural and temporal analyses including collaboration networks, author co-citation networks, and document co-citation networks with support of hybrid node types such as terms, institutions, and countries, and hybrid link types such as co-citation, co-occurrence, and directed citing links.
CiteWiz [12]	C	-	-	Visualization of citation networks using causality visualization techniques, interactive timelines, and concept maps.
Proposed tool ReViz	C	✓	✓	Tool especially designed for systematic reviews, and running stand-alone or integrated in Parsifal for visualizing citation graphs constructed from a set of paper documents. Support of various approaches for simplifying citation graphs.

* C: Focus on Citation Networks, B: General Bibliometric Network Visualizations

- **ReViz: lightweight software for generating citation graphs**
 - Optional: Integration into survey tool "Parsifal"

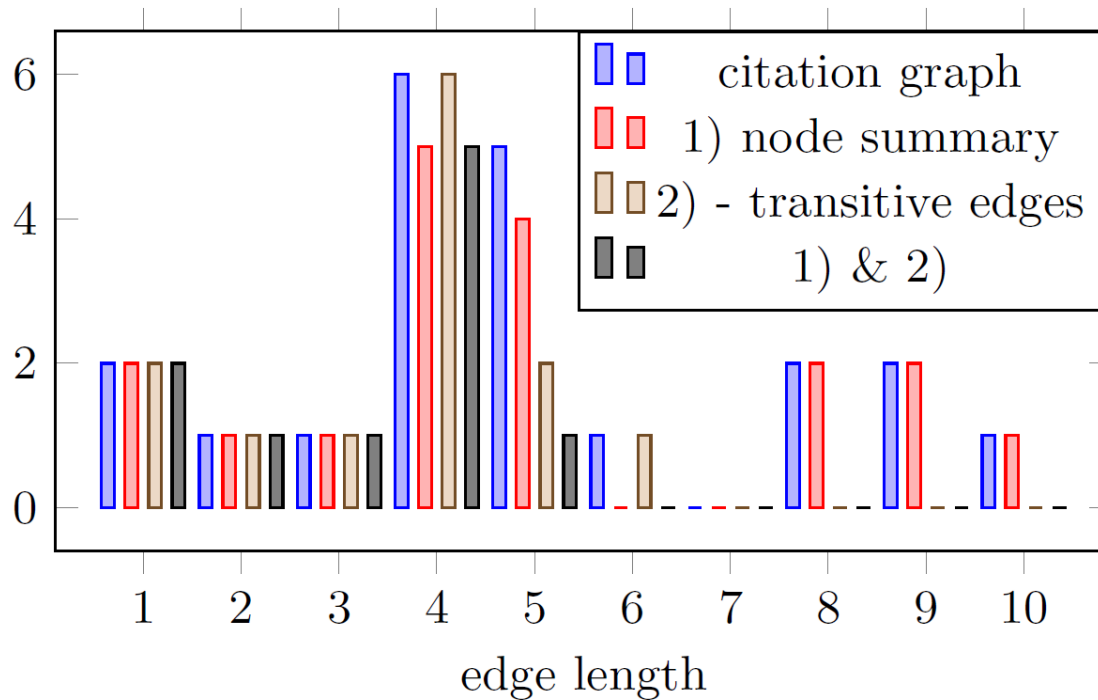
Evaluation

- Evaluation according to **objective criteria**
 - # nodes
 - # edges
 - # edges/nodes
 - # hiding nodes
 - # hidden edges
 - # overlaps
- **User evaluation**
 - Survey & analysis of results

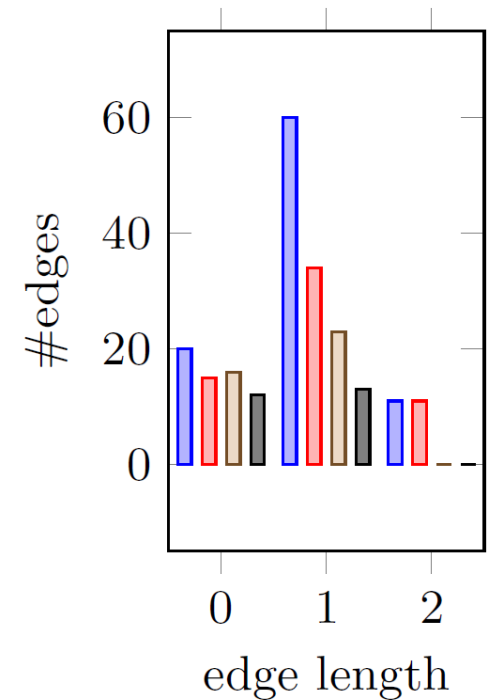
Evaluation of Graph Properties 1/2

		Nodes	Edges	Edges per Node
Graph 1 [17]	Normal Graph	12	21	3.5
	1. Node Summary	11 (-8%)	18 (-14%)	3.27 (-7%)
	2. Transitivity	12 (0%)	13 (-38%)	2.17 (-38%)
	1. & 2.	11 (-8%)	10 (-52%)	1.81 (-48%)
Graph 2 [2]	Normal Graph	22	92	8.36
	1. Node Summary	18 (-18%)	60 (-35%)	6.67 (-20%)
	2. Transitivity	22 (0%)	40 (-57%)	3.64 (-56%)
	1. & 2.	18 (-18%)	25 (-73%)	2.78 (-67%)
		Hiding Nodes	Hidden Edges	Overlaps
Graph 1 [17]	Normal Graph	4	8	~42
	1. Node Summary	4 (0%)	10 (+25%)	33 (-21%)
	2. Transitivity	3 (-25%)	2 (-75%)	7 (-83%)
	1. & 2.	3 (-25%)	2 (-75%)	4 (-90%)
Graph 2	Normal Graph [2]	18	~77	~220
	1. Node Summary	14 (-22%)	~47 (-39%)	~90 (-59%)
	2. Transitivity	18 (0%)	~32 (-58%)	~44 (-80%)
	1. & 2.	13 (-28%)	~19 (-75%)	10 (-95%)

Evaluation of Graph Properties 2/2



(a) Graph 1

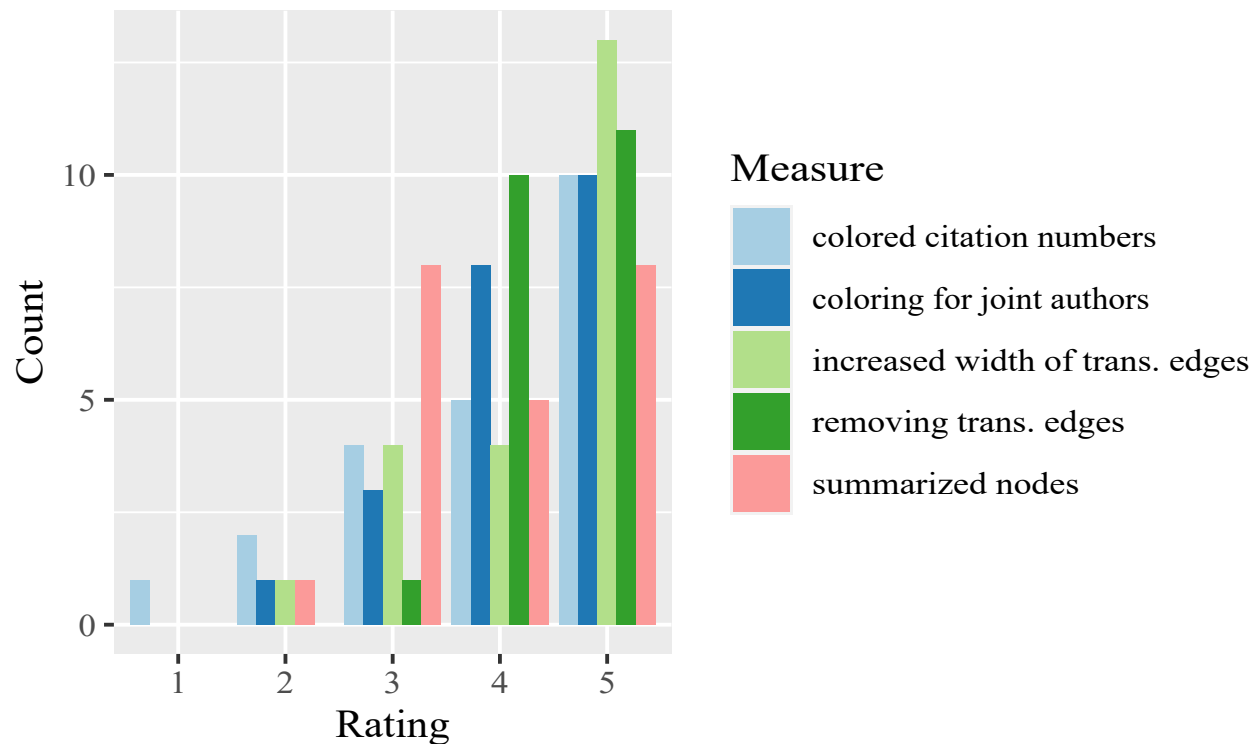


(b) Graph 2

- edge length on the x-axis indicates how many layers (i.e., years) an edge passes over

User Evaluation

- 22 participants



- rating 1: not useful at all ...
- rating 5: very useful

User Evaluation

- 95% confidence intervals for improving the clarity, layout of the graphs and comprehensibility using the different functionalities

Functionality	Clarity	Layout	Comprehensibility
Node Summary	[4.00, 4.45]	[3.36, 4.10]	[2.99, 3.74]
Removing Transitive Edges	[4.50, 5.00]	[4.00, 4.64]	[3.78, 4.18]
Edge thickness of transitive edges	[3.49, 4.24]	[3.27, 4.10]	[3.73, 4.45]
Colored Number of Citations	[3.21, 4.25]	[3.38, 4.35]	[3.42, 4.22]

- rating 1: not useful at all ...
- rating 5: very useful

Summary & Conclusions

- ReViz tool: Visualizations for survey papers
 - (optional) integration into survey tool Parsifal
 - Flow chart for selection process of publications
 - Citation graphs with variants
 - simple
 - node summaries
 - removing transitive edges
 - number of (in-) direct citations in node labels
 - coloring publications with common authors
- Evaluation including user evaluation
 - variants reduce # nodes, # edges, # hiding nodes, # hidden edges, # overlaps
 - positively rated by users