Seminar Software Composition: Project P6

How are Software Visualizations Evaluated?

Presented by: Lukas Imstepf

lukas.imstepf@unifr.ch

Department of Informatics
University of Fribourg

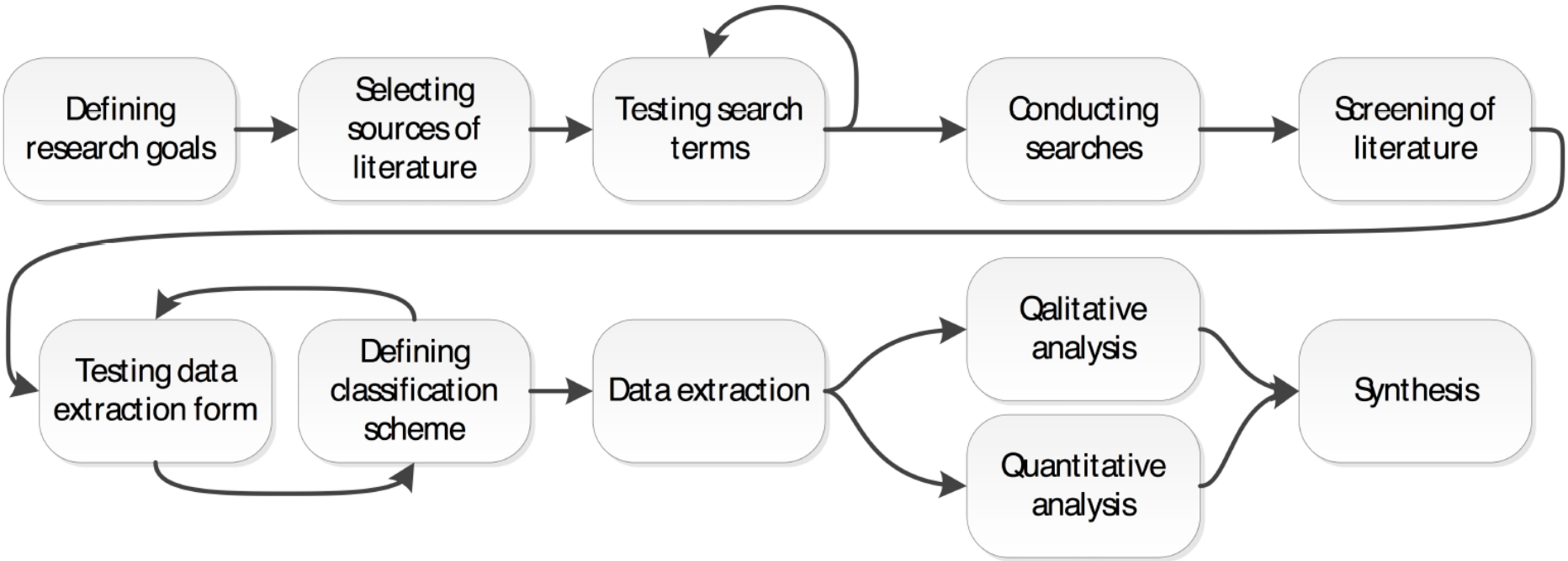
June 27, 2017

MASTER IN COMPUTER SCIENCE

Research Method

Systematic Literature Review

Research Method



- Following Kitchenham's guidlines for systematic literature reviews in software engineering.

B. Kitchenham (2004): *"Procedures for Performing Systematic Reviews"*

Search Strategy and Data Sources

Research Method

1. Search for:

```
"software visualization" OR
"software visualisation"
```

Search Strategy and Data Sources

Research Method

1. Search for:

   ```
   "software visualization" OR
   "software visualisation"
   ```

2. in three scientific online databases:
   - ACM Digital Library
   - IEEE Xplore DL
   - ScienceDirect

Search Strategy and Data Sources

Research Method

1. Search for:

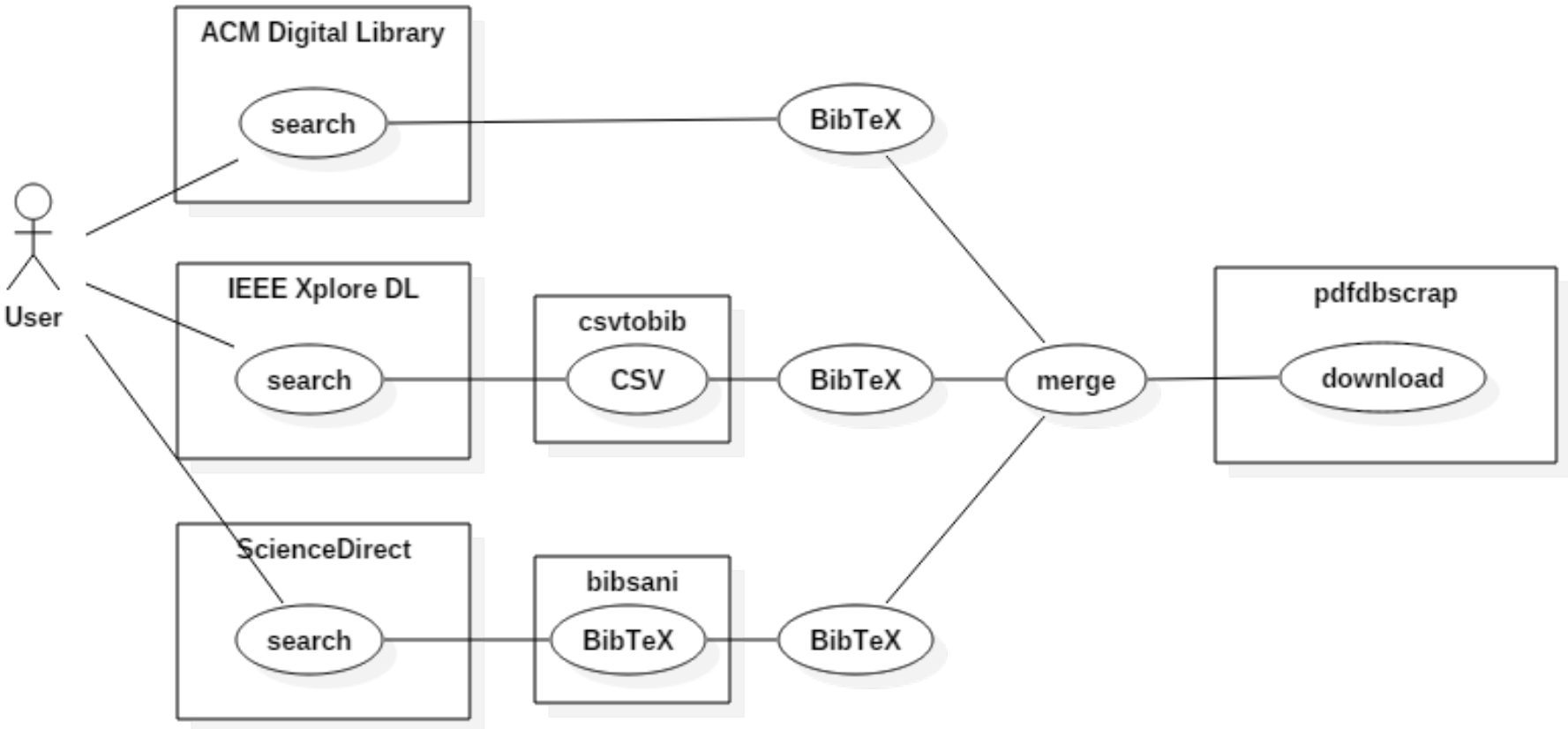
   ```
   "software visualization" OR
   "software visualisation"
   ```

2. in three scientific online databases:
   - ACM Digital Library
   - IEEE Xplore DL
   - ScienceDirect

3. Download the search results as BibTeX/CSV files, converting CSV to BibTeX with *bibsani*

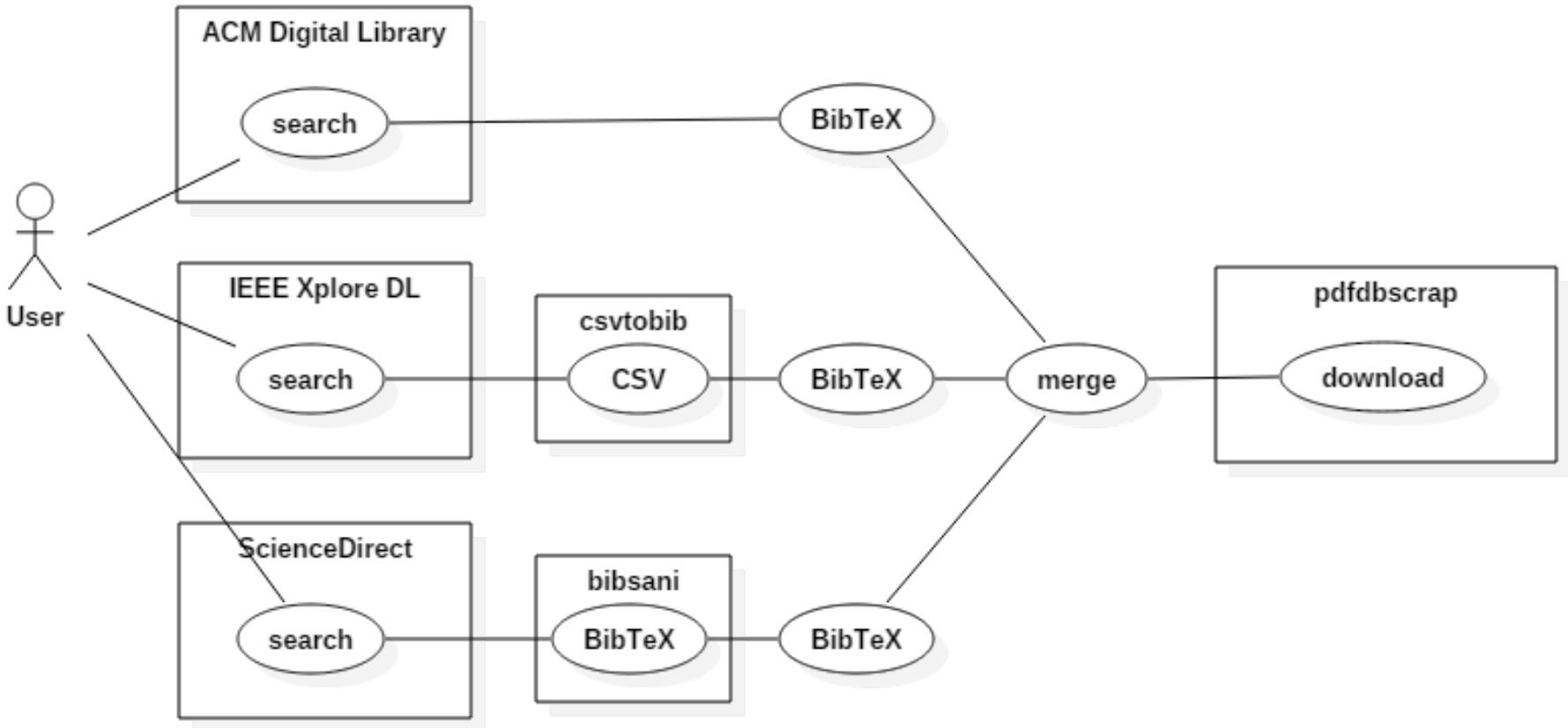
Data Scraping and Early Exclusion Criteria

Research Method



Data Scraping and Early Exclusion Criteria

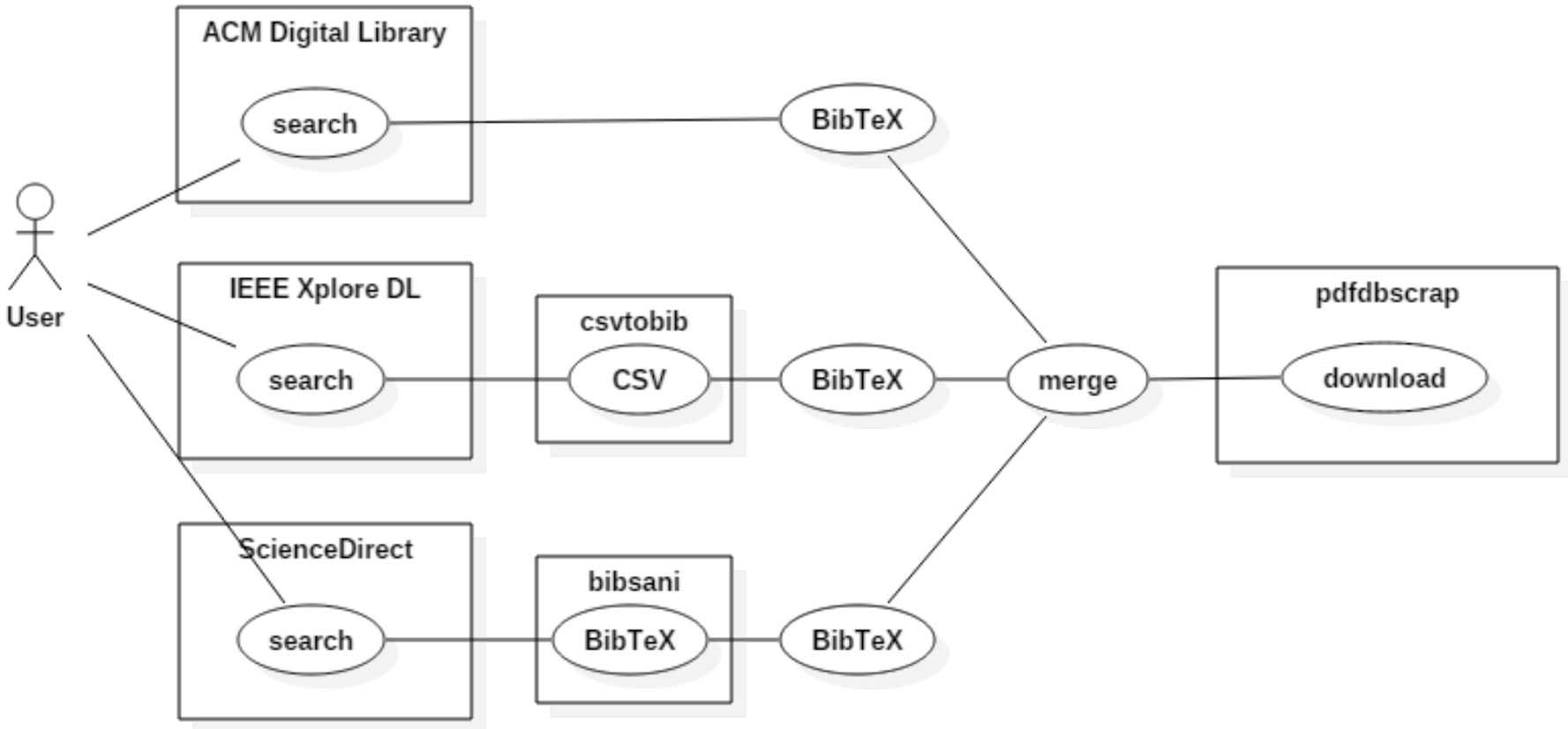
Research Method



- 1289 search results

Data Scraping and Early Exclusion Criteria

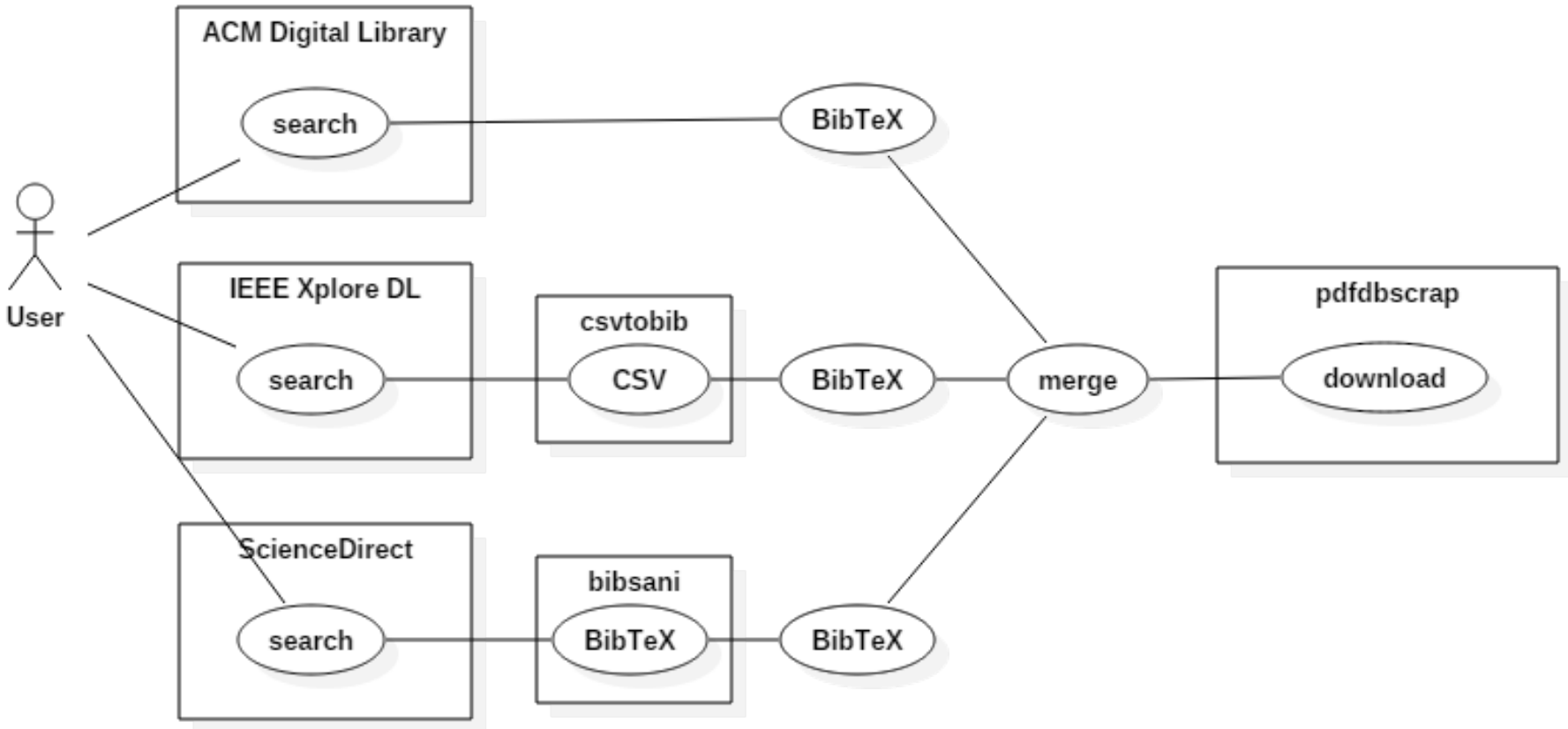
Research Method



- 1289 search results
- Exclusion criteria: data sanitization (incomplete entries, duplicates: -100)

Data Scraping and Early Exclusion Criteria

Research Method



- 1289 search results
- Exclusion criteria: data sanitization (incomplete entries, duplicates: -100)
- Exclusion criteria: scraping failures (-66, $+7$ manual downloads)

Exclusion Criteria

Research Method

- Subtotal: 1130 scrapped PDF files

Exclusion Criteria

Research Method

- Subtotal: 1130 scrapped PDF files

Further exclusion criteria:

1. Fewer than five pages $(-279, \text{subtotal} = 851)$

Exclusion Criteria

Research Method

- Subtotal: 1130 scrapped PDF files

Further exclusion criteria:

1. Fewer than five pages (-279, subtotal $= 851$)
2. InfoVis (medical/geographical) papers (-318, subtotal $= 533$)

Exclusion Criteria

Research Method

- Subtotal: 1130 scrapped PDF files

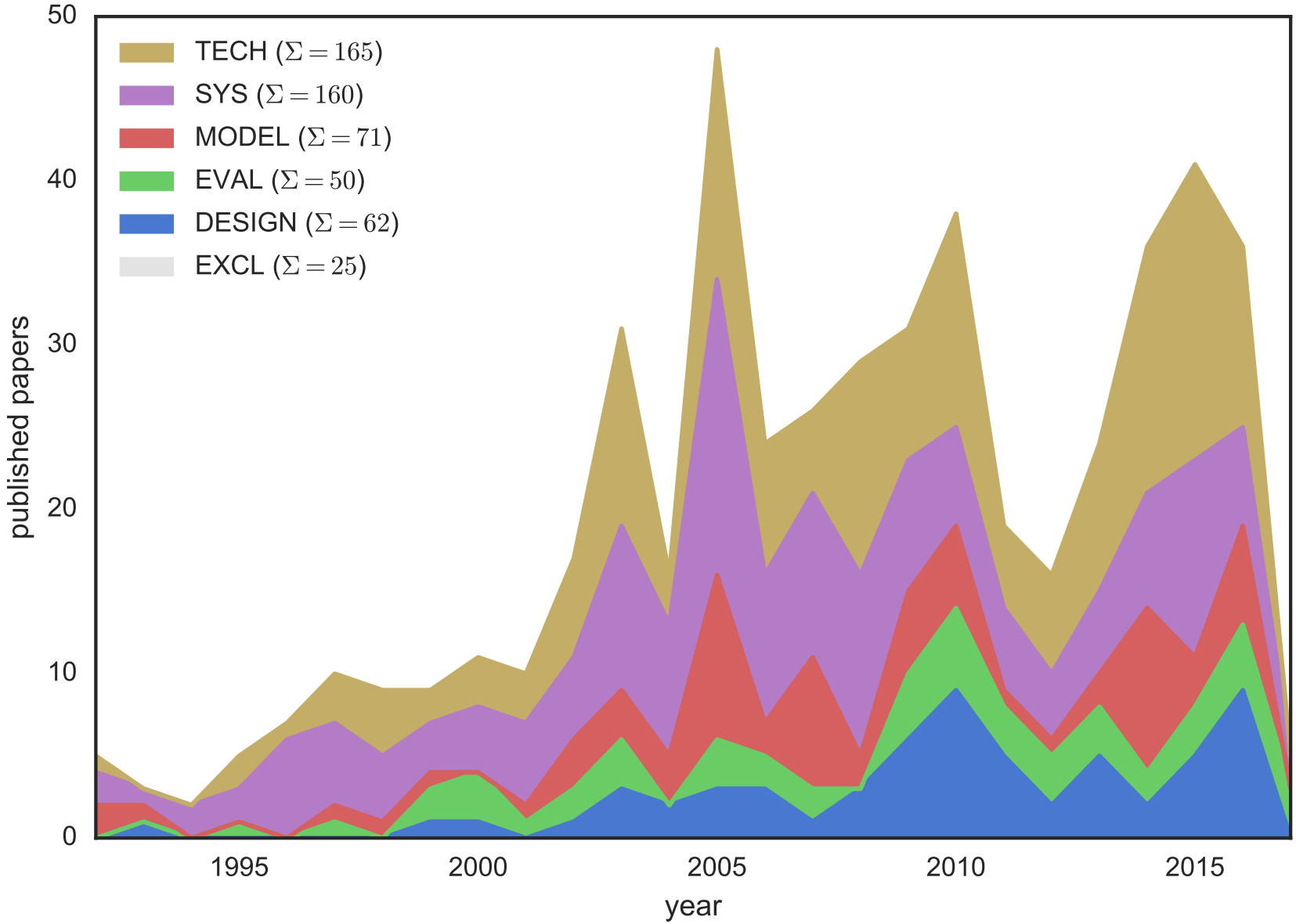
Further exclusion criteria:

1. Fewer than five pages (−279, subtotal = 851)
2. InfoVis (medical/geographical) papers (−318, subtotal = 533)
3. Exclusion by paper type:
   - 3.1 Technique papers (novel algorithms)
   - 3.2 Design study papers (particular domain problems)
   - 3.3 Systems papers (architectural choices)
   - 3.4 Evaluation papers
   - 3.5 Model papers (taxonomy, formalisms, commentary)

Paper types by Munzner (2008): *"Process and Pitfalls in Writing Information Visualization Research Papers"*.

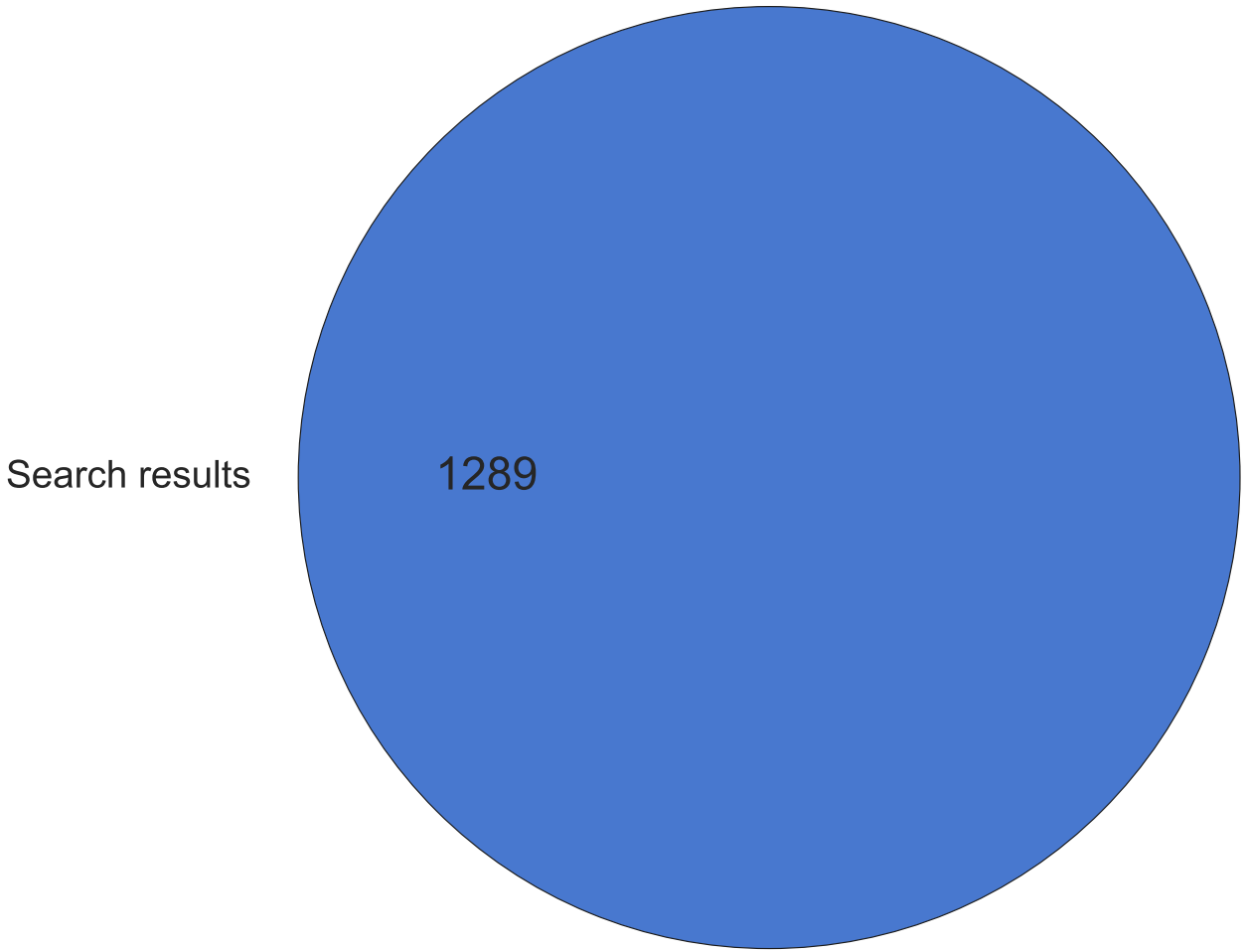
Exclusion Criteria: Paper Type

SoftVis papers from 1992 to 2017



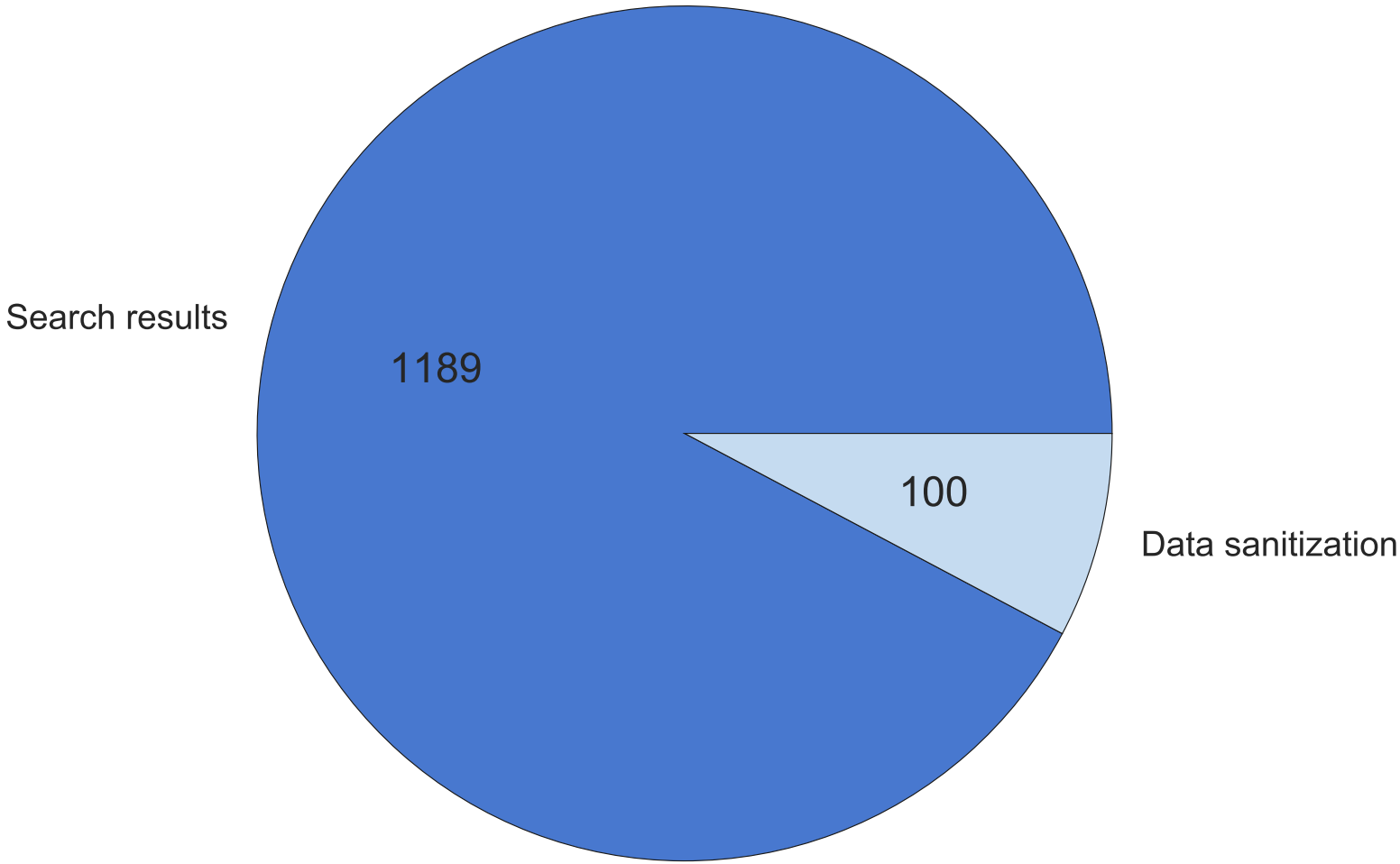
Exclusion Criteria: Recap

Research Method



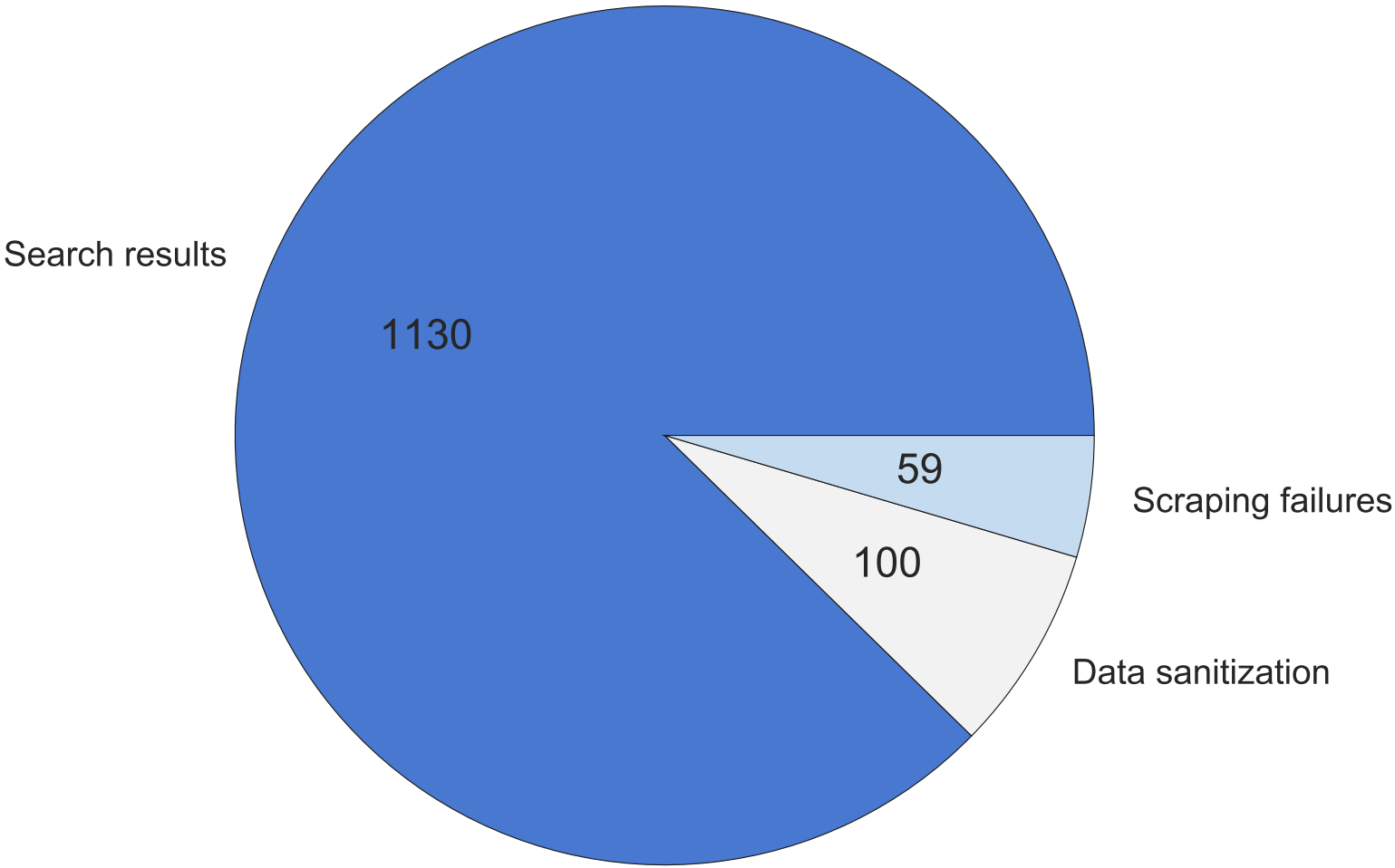
Exclusion Criteria: Recap

Research Method



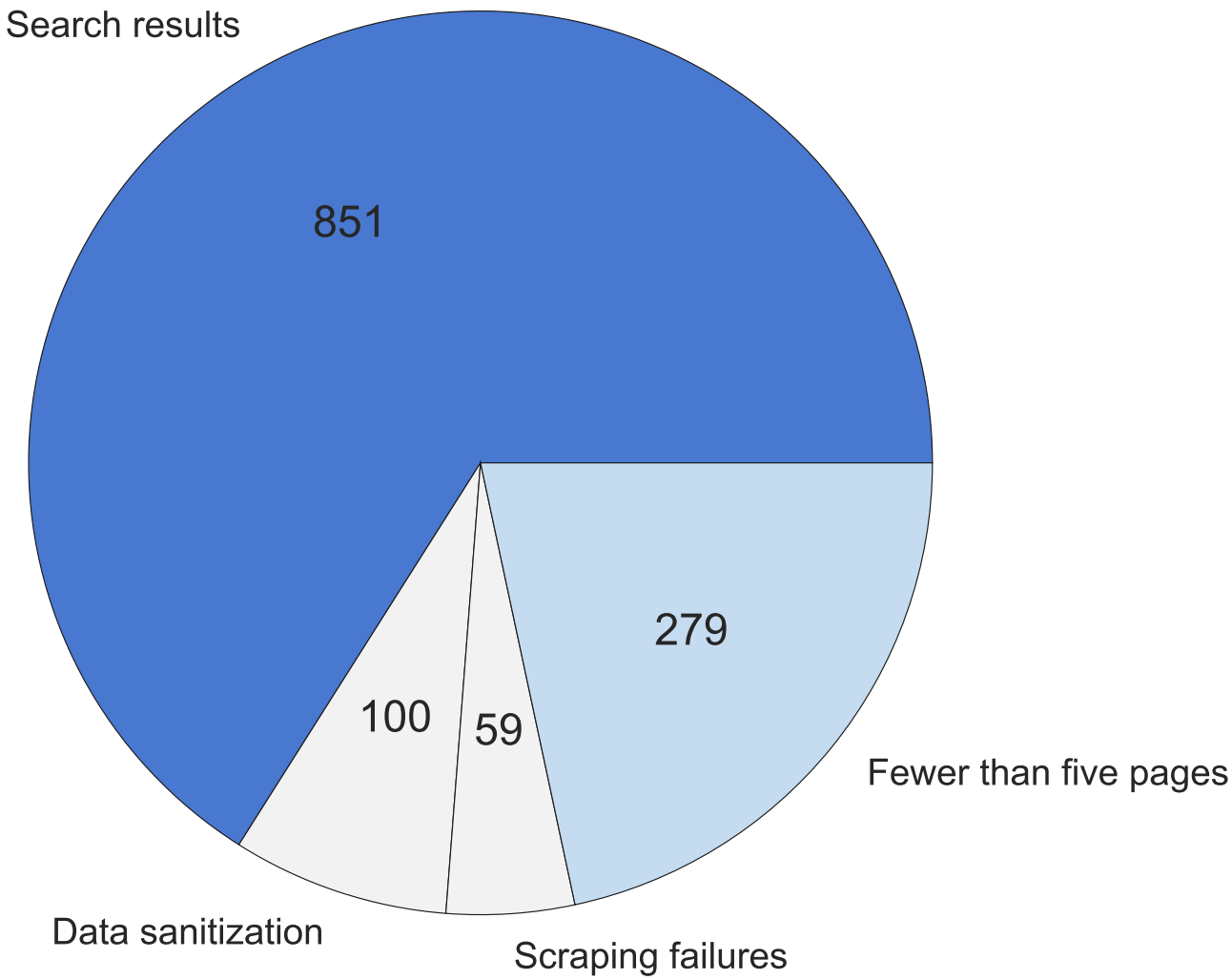
Exclusion Criteria: Recap

Research Method



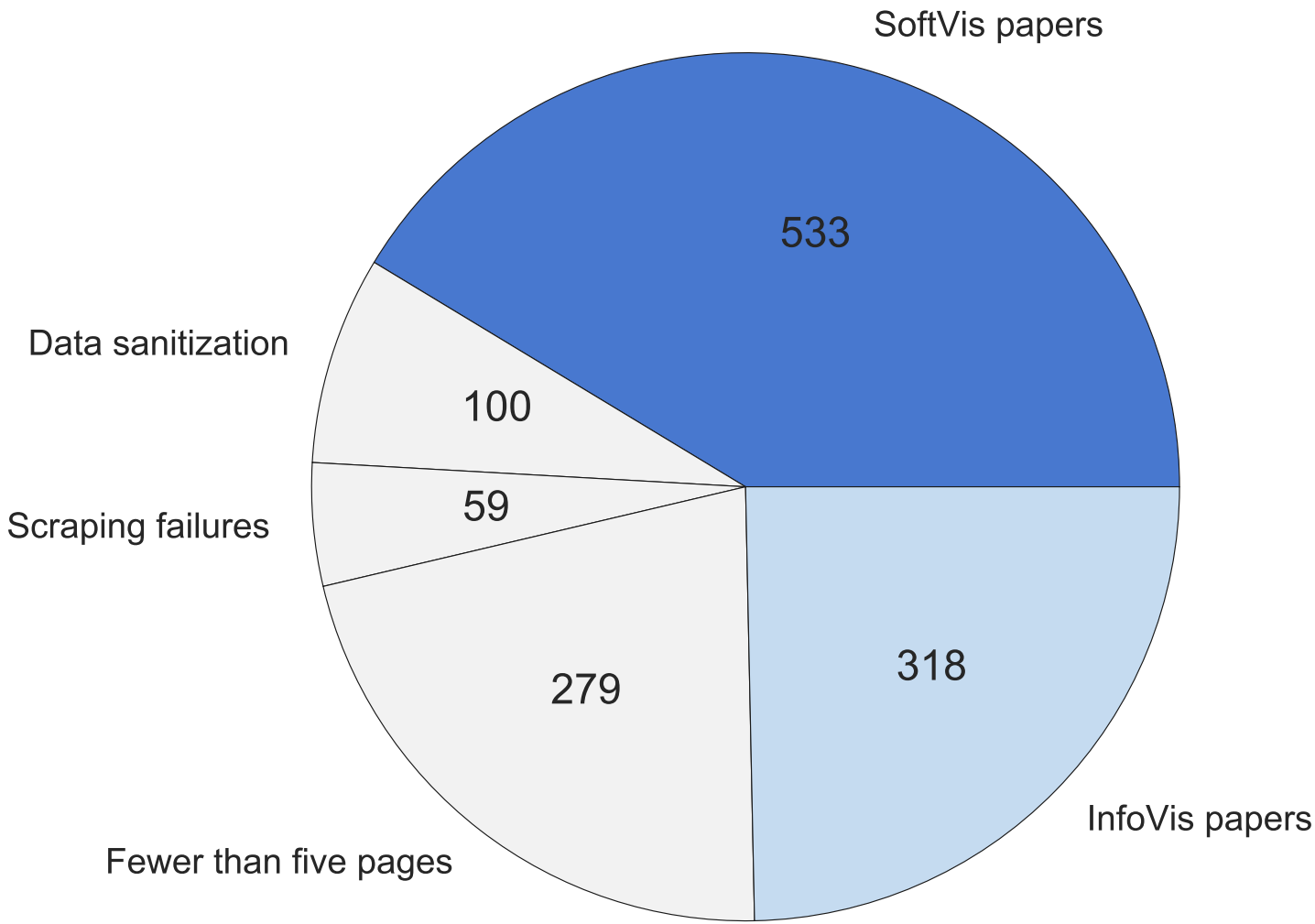
Exclusion Criteria: Recap

Research Method



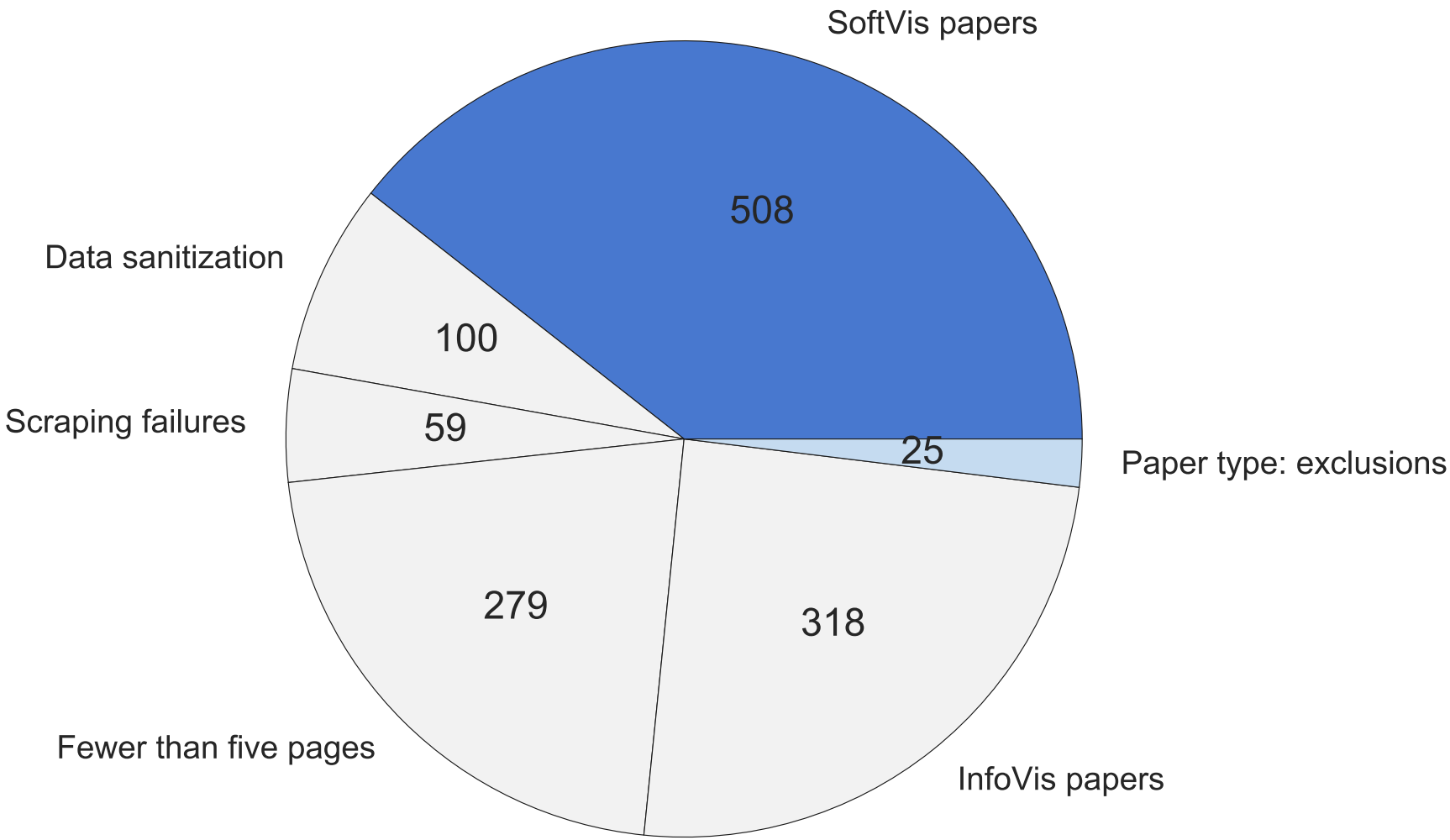
Exclusion Criteria: Recap

Research Method



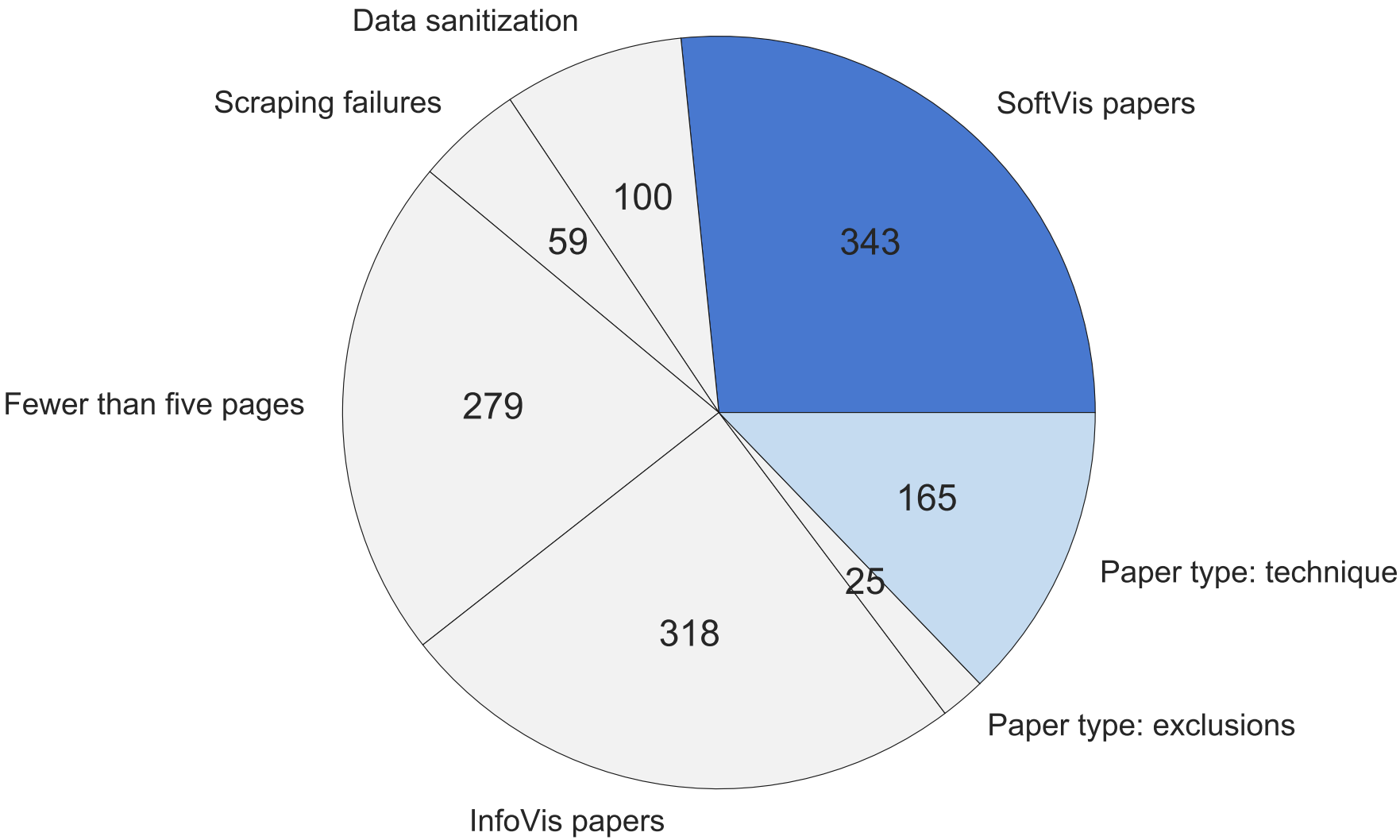
Exclusion Criteria: Recap

Research Method



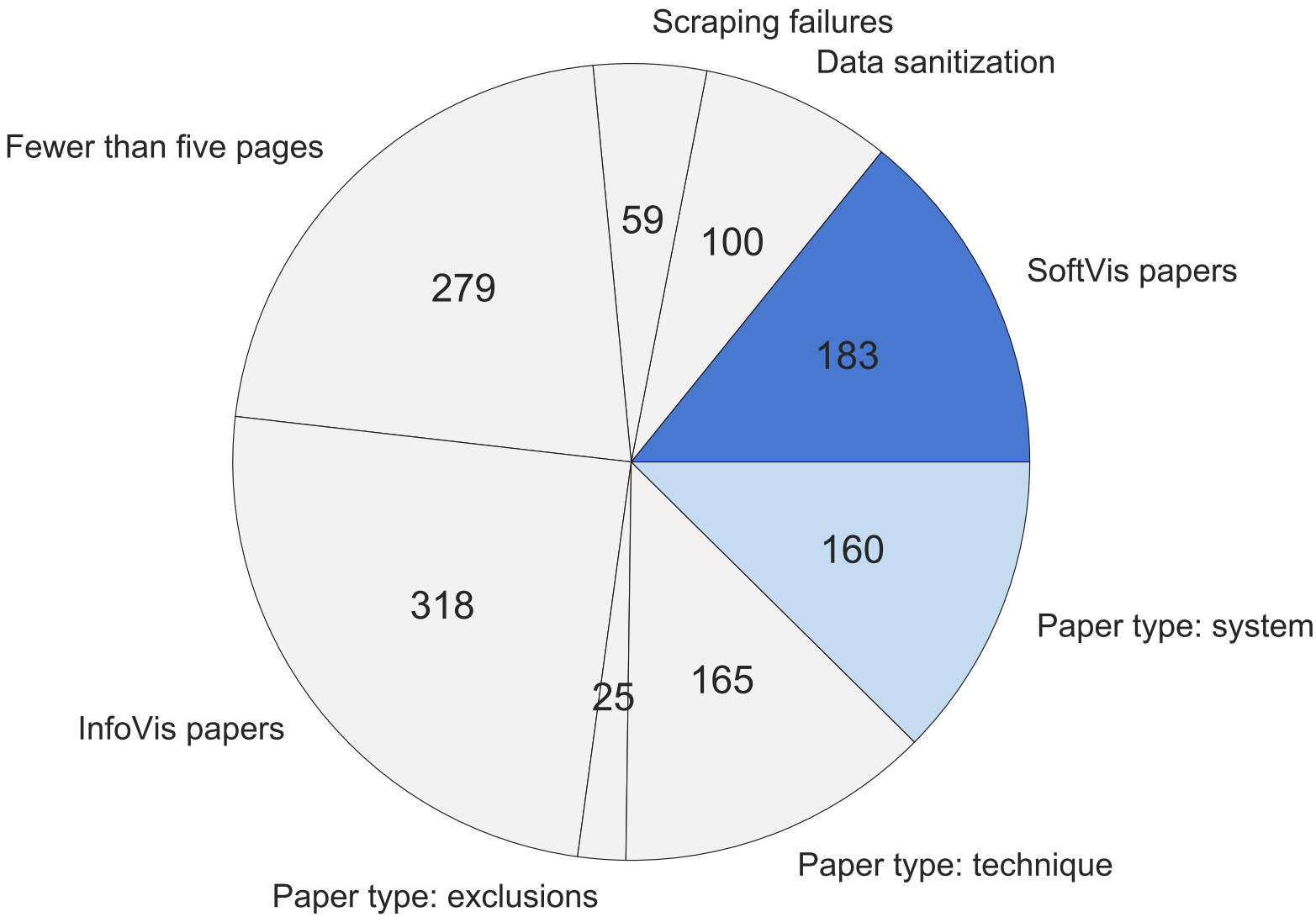
Exclusion Criteria: Recap

Research Method



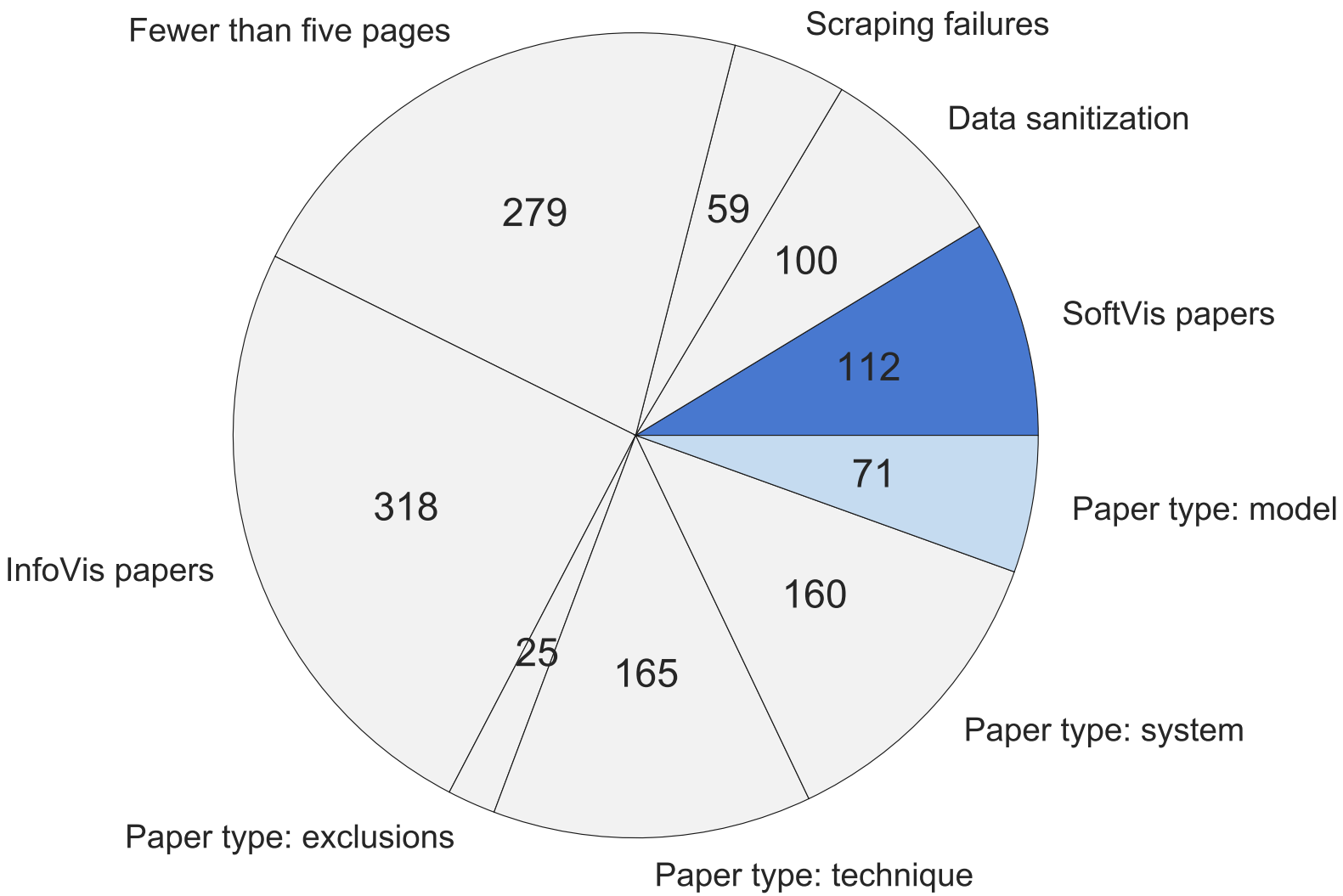
Exclusion Criteria: Recap

Research Method



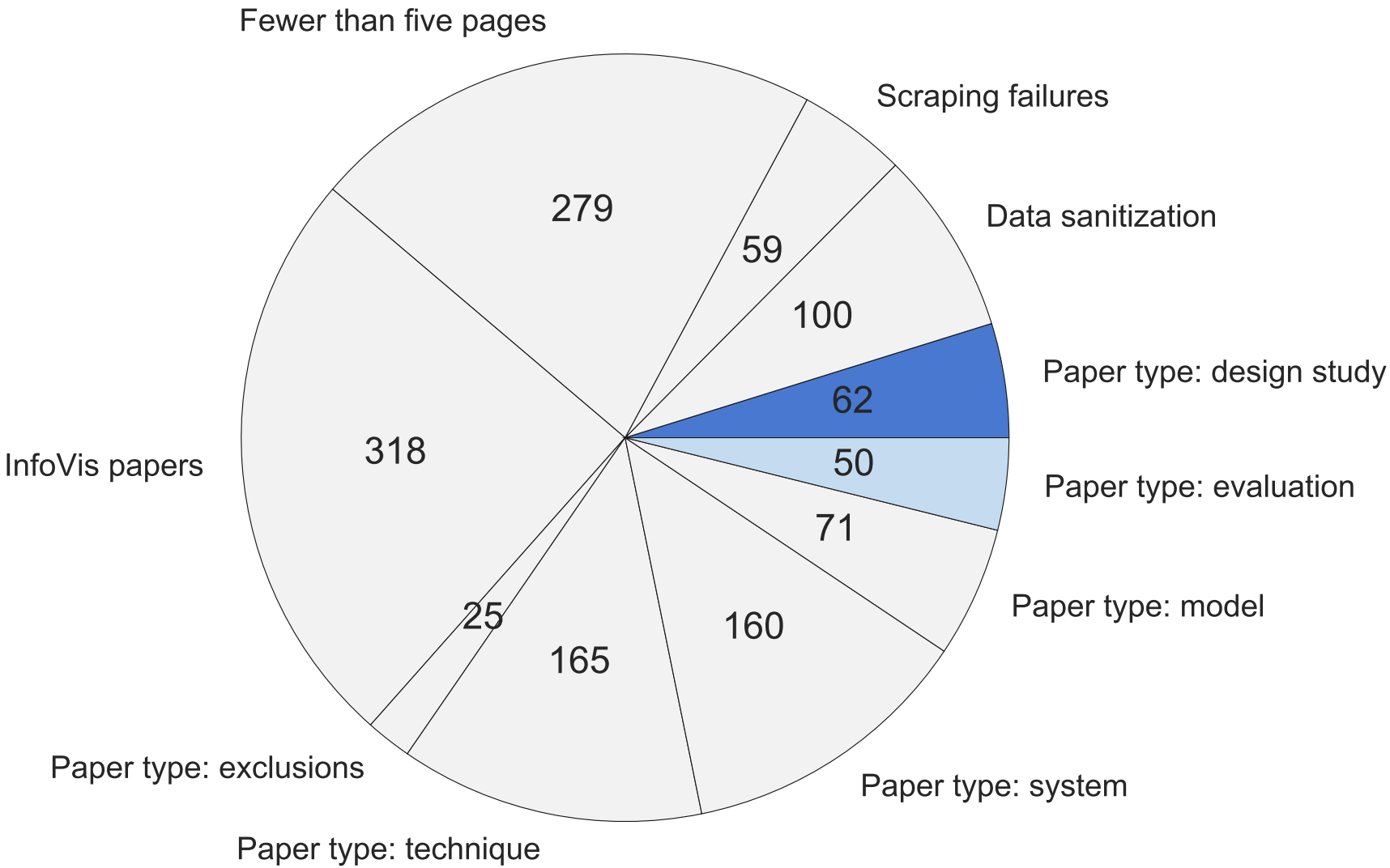
Exclusion Criteria: Recap

Research Method



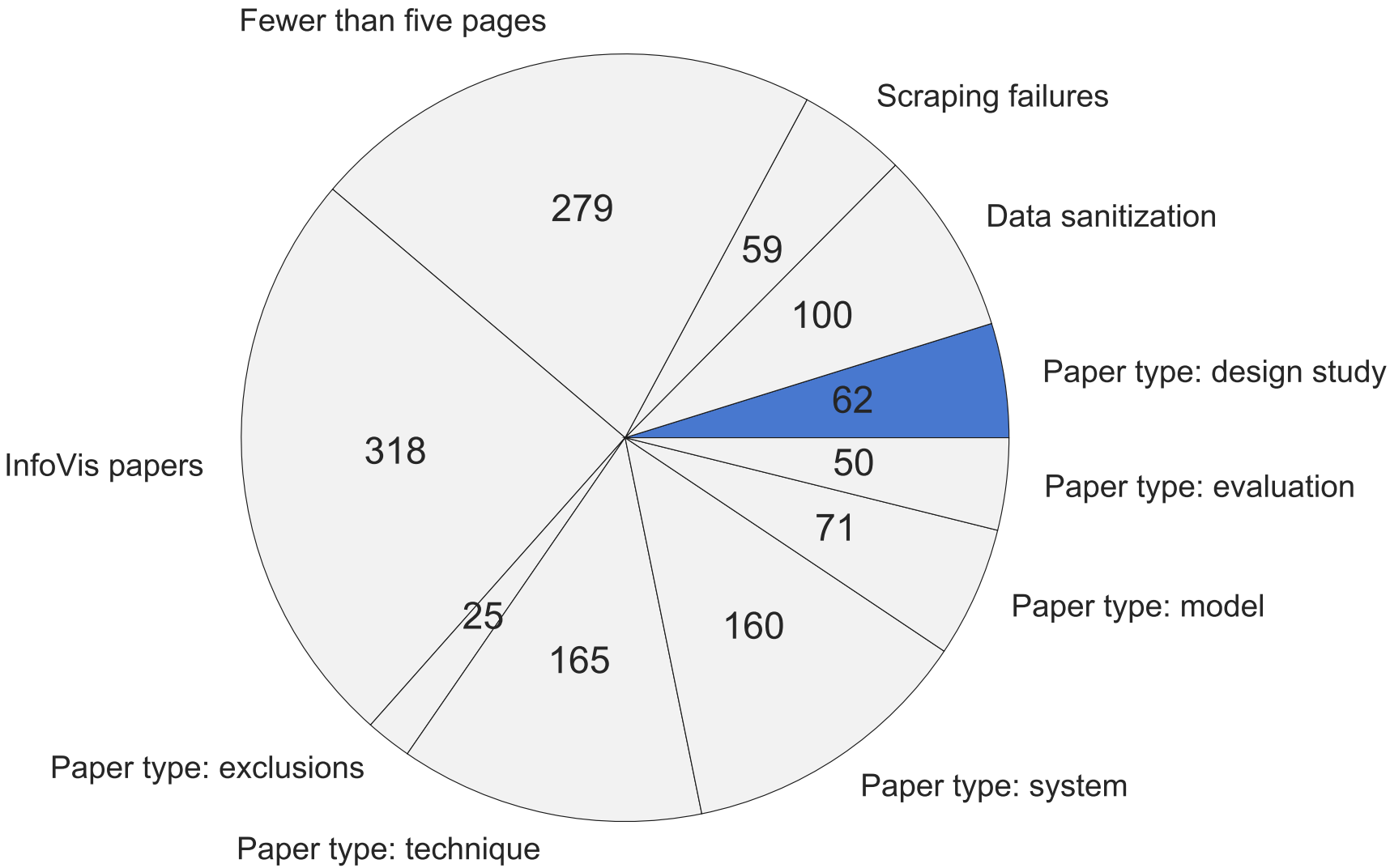
Exclusion Criteria: Recap

Research Method



Exclusion Criteria: Recap

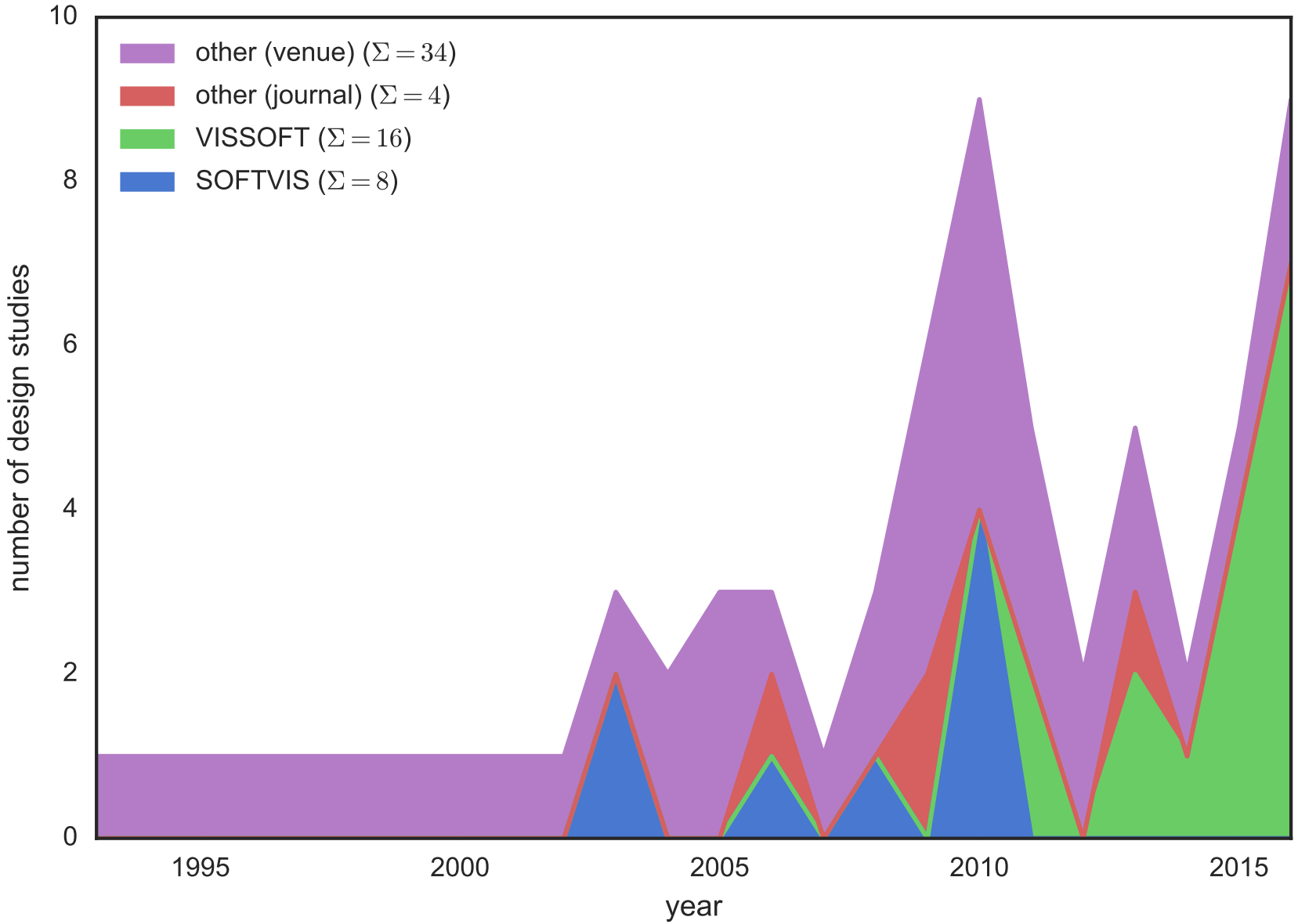
Research Method



62 Design studies

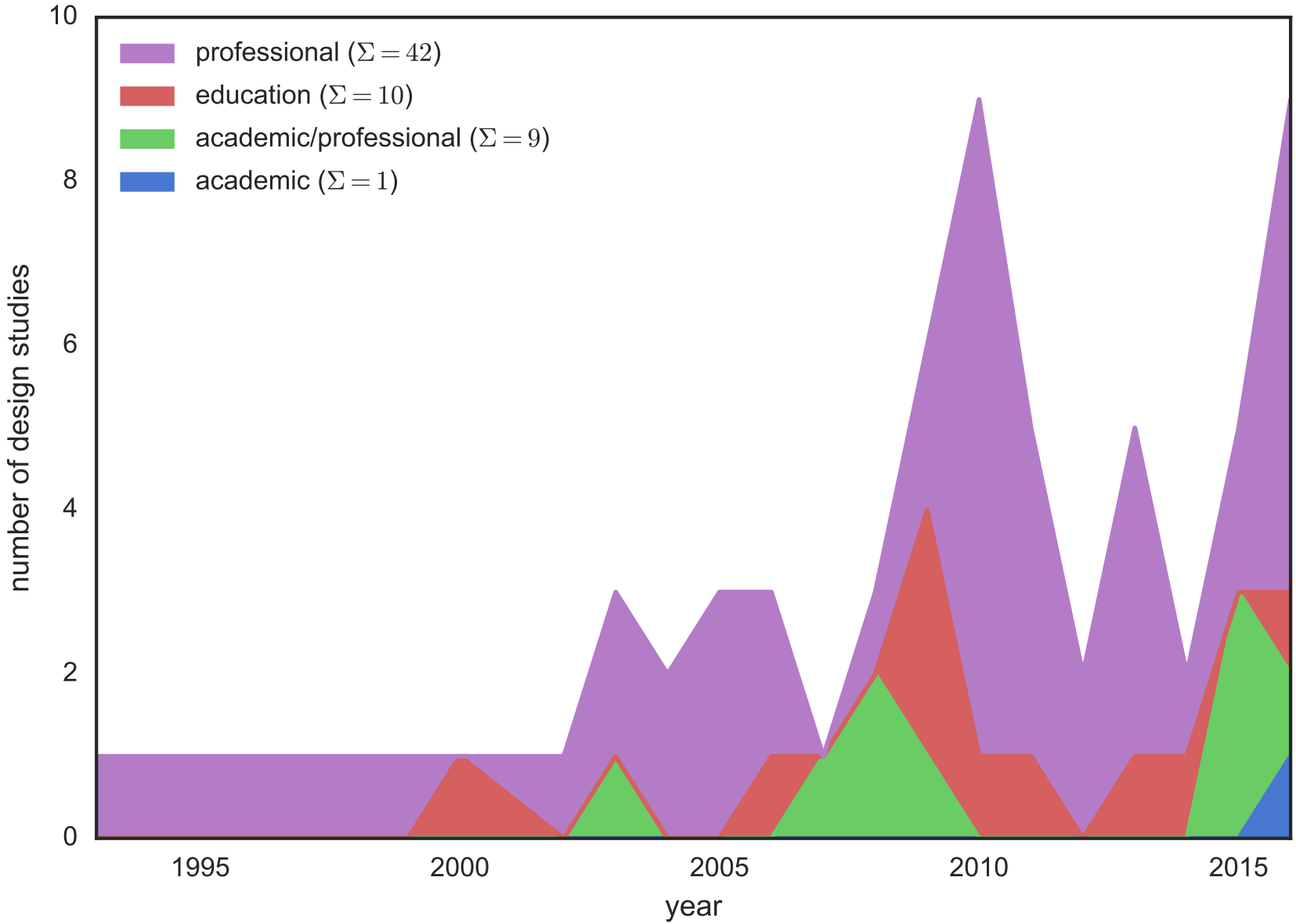
Venues

Design studies, $N = 62$



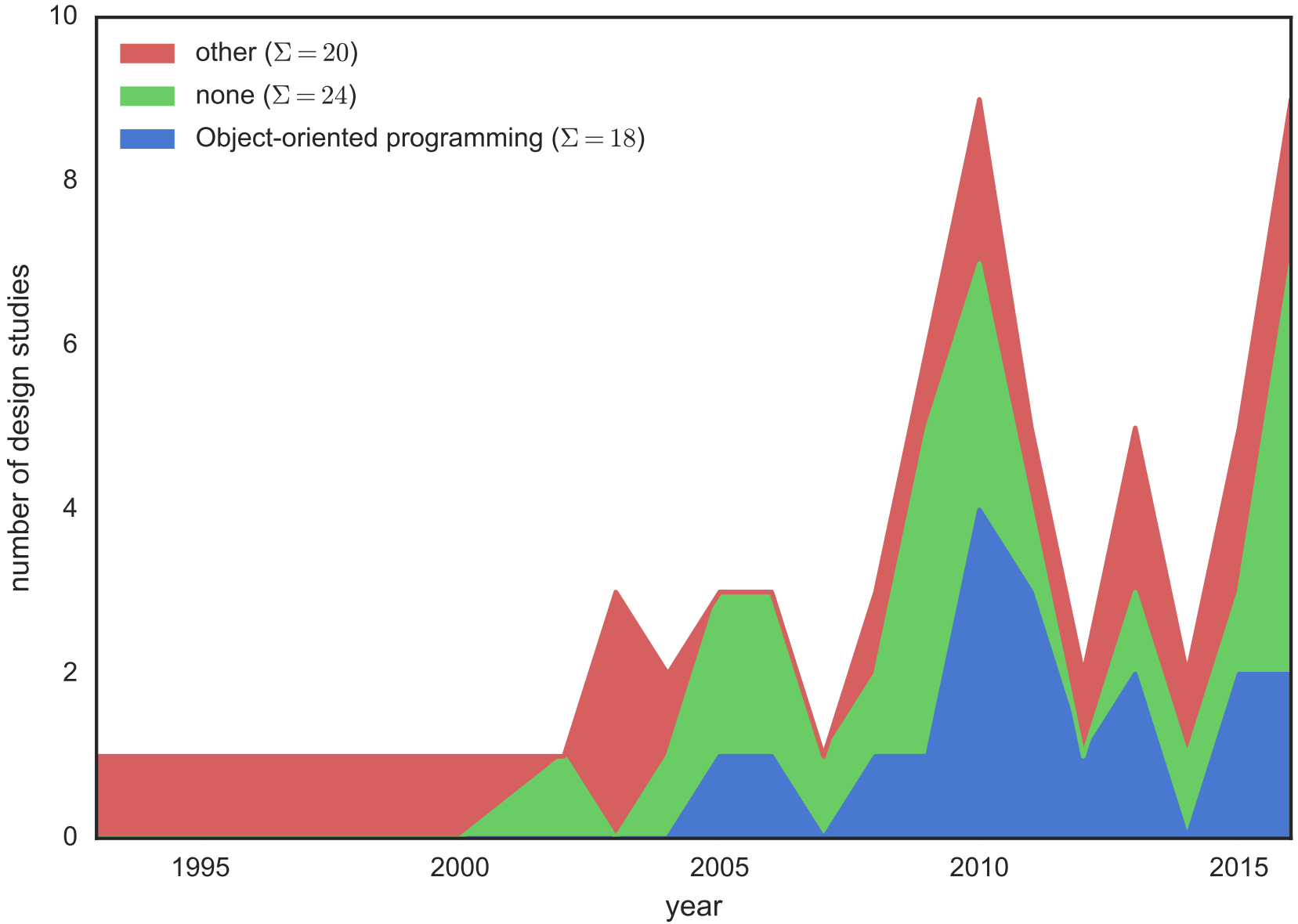
Target audience

Design studies, $N = 62$



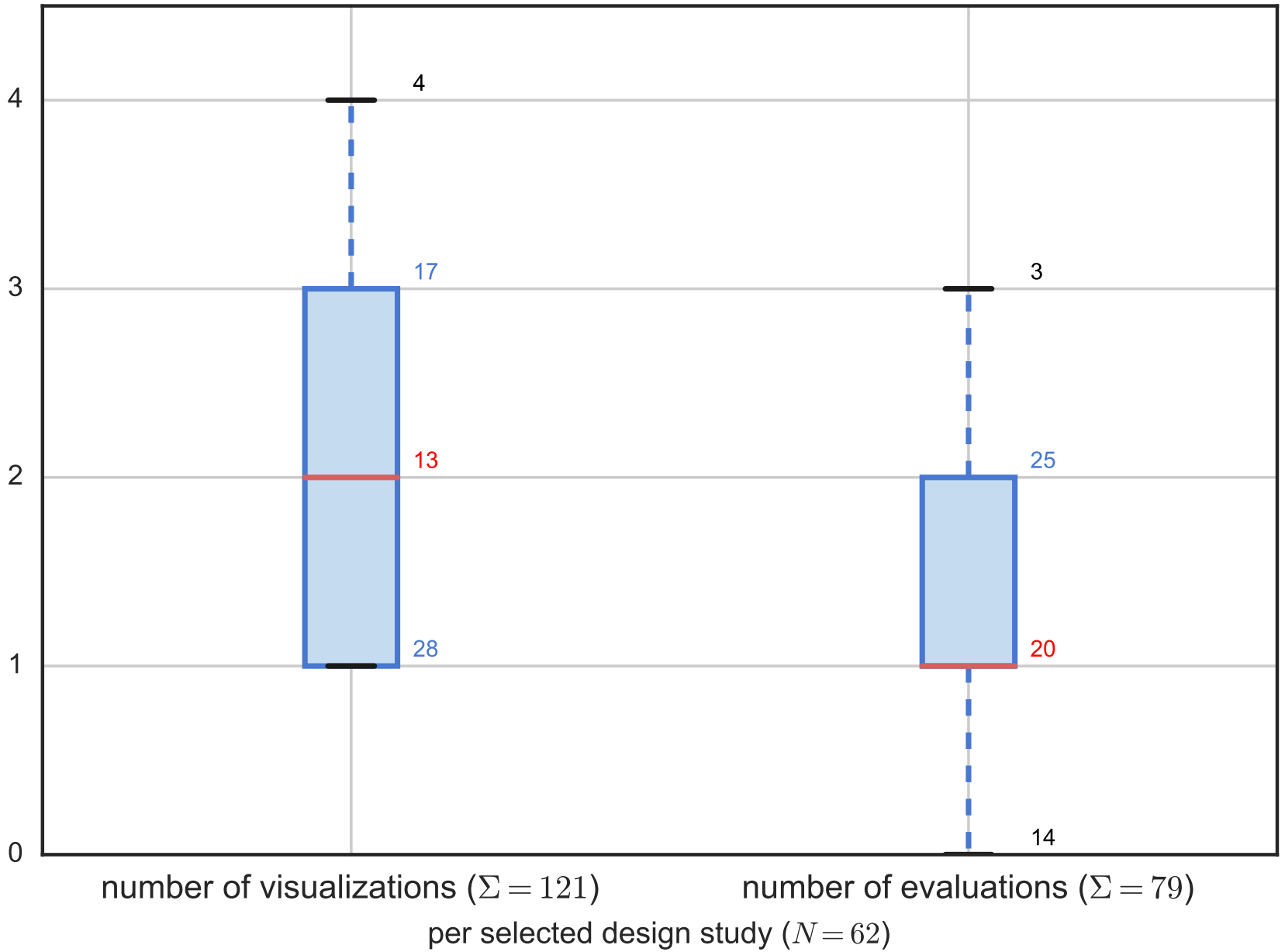
Programming Paradigms

Design studies, $N = 62$



Visualizations and evaluations

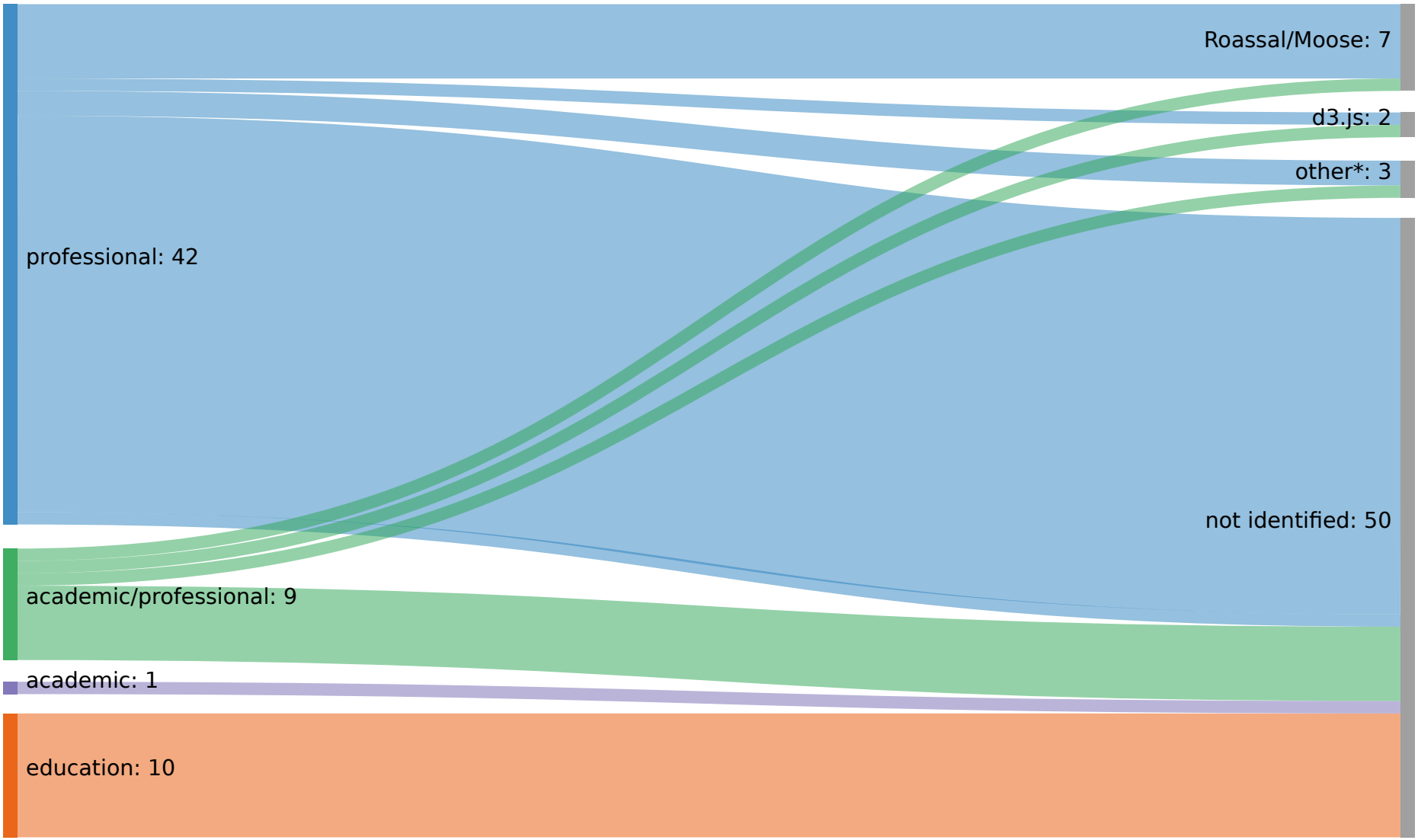
Overview



121 Visualizations of 62 design studies

Usage of visualization frameworks

Design study papers, $N = 62$



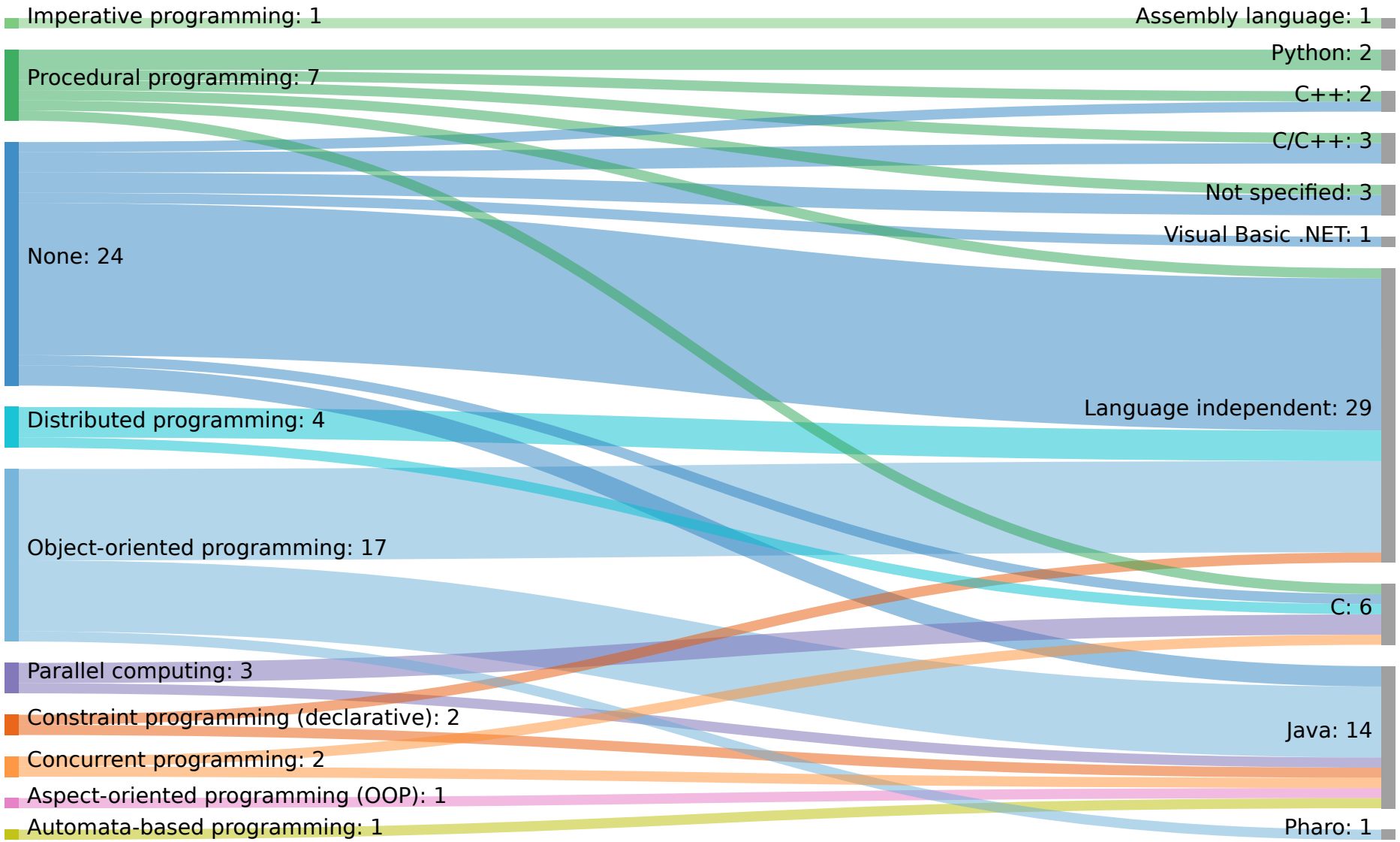
*other visualization frameworks: *Graphplace*, *Flatland*, *Sovis*

Presentation of a new visualization tool?

Design study papers, $N = 62$

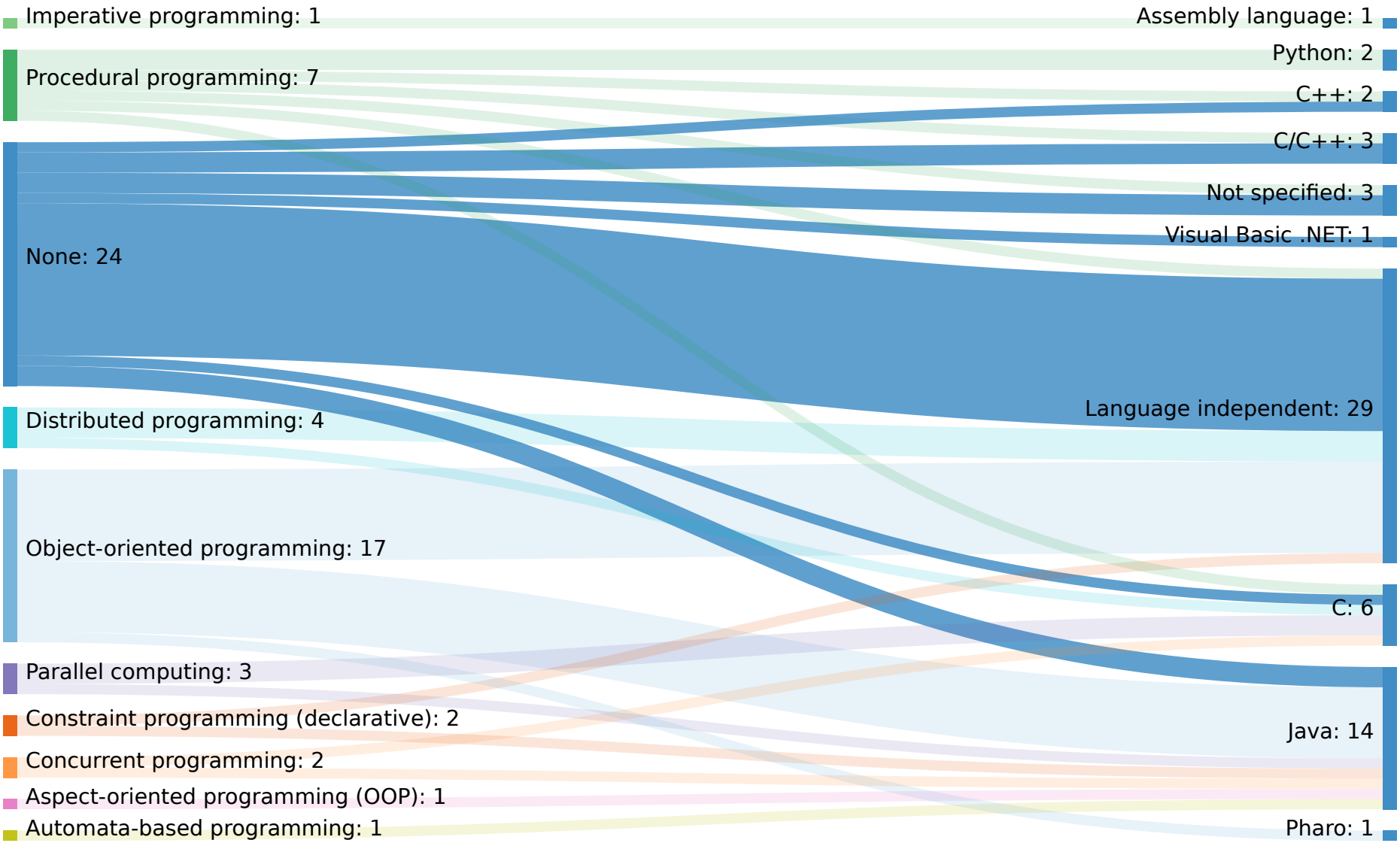
Who needs visualizations, and why?

Programming paradigms vs. programming languages



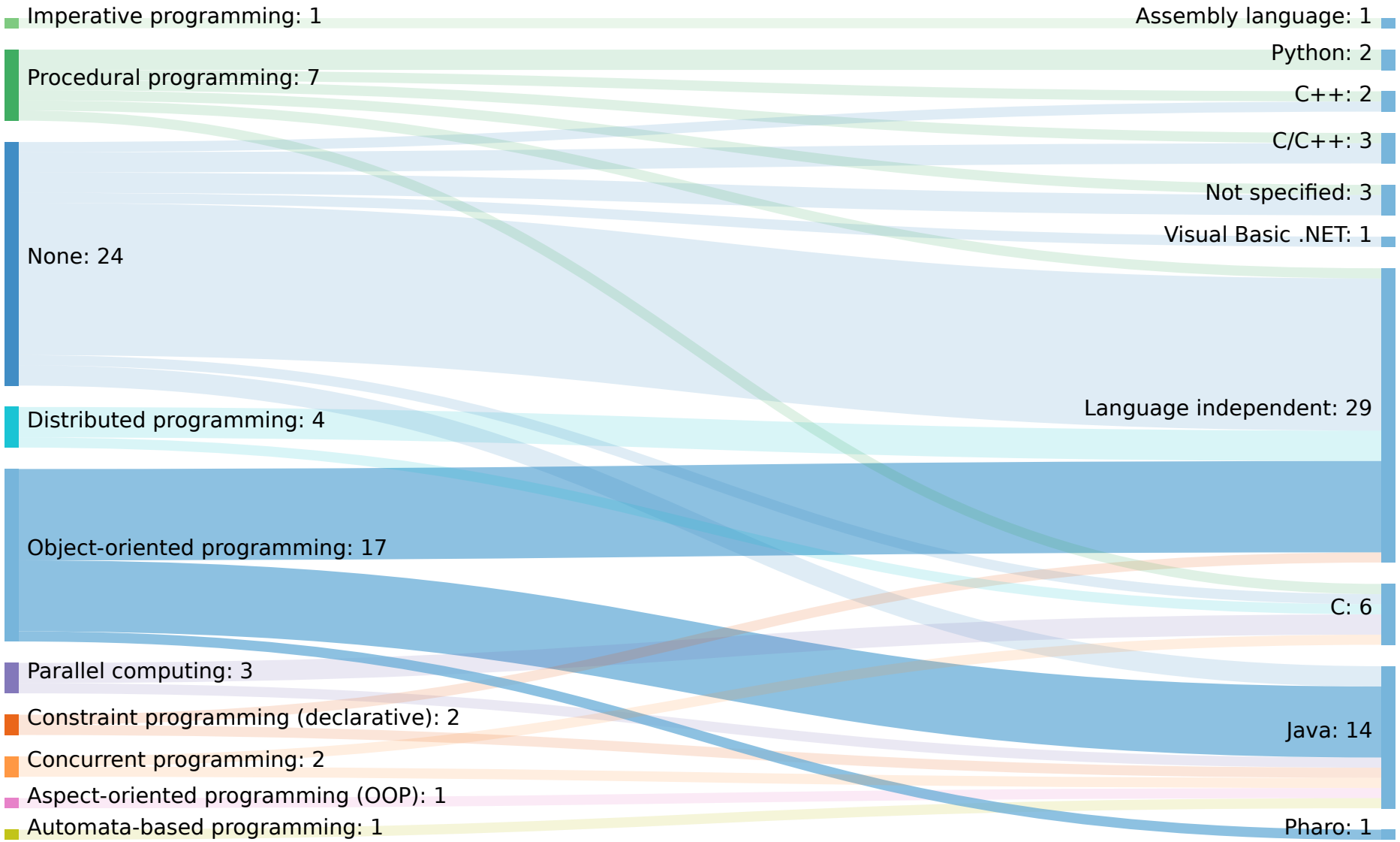
Who needs visualizations, and why?

Programming paradigms vs. programming languages



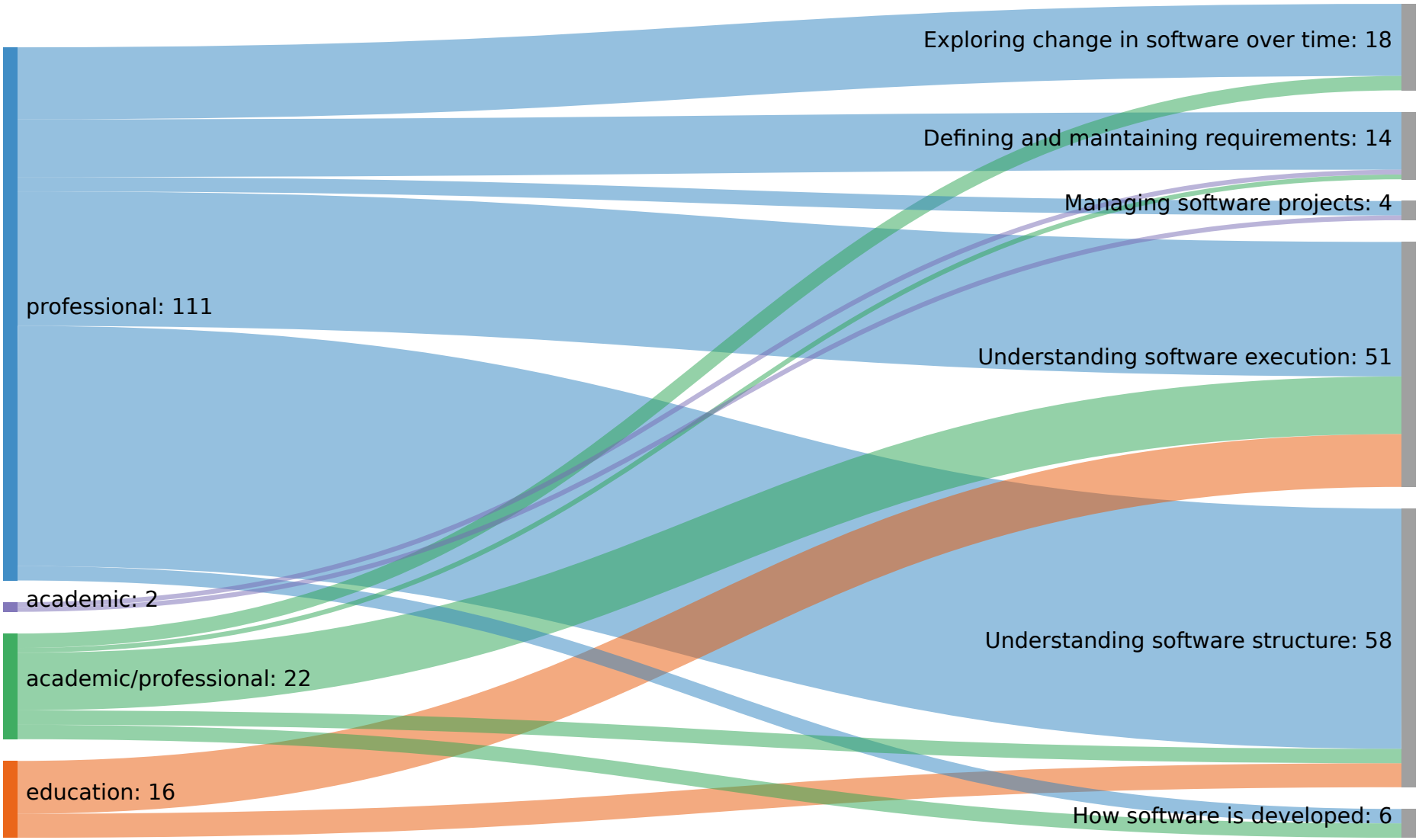
Who needs visualizations, and why?

Programming paradigms vs. programming languages



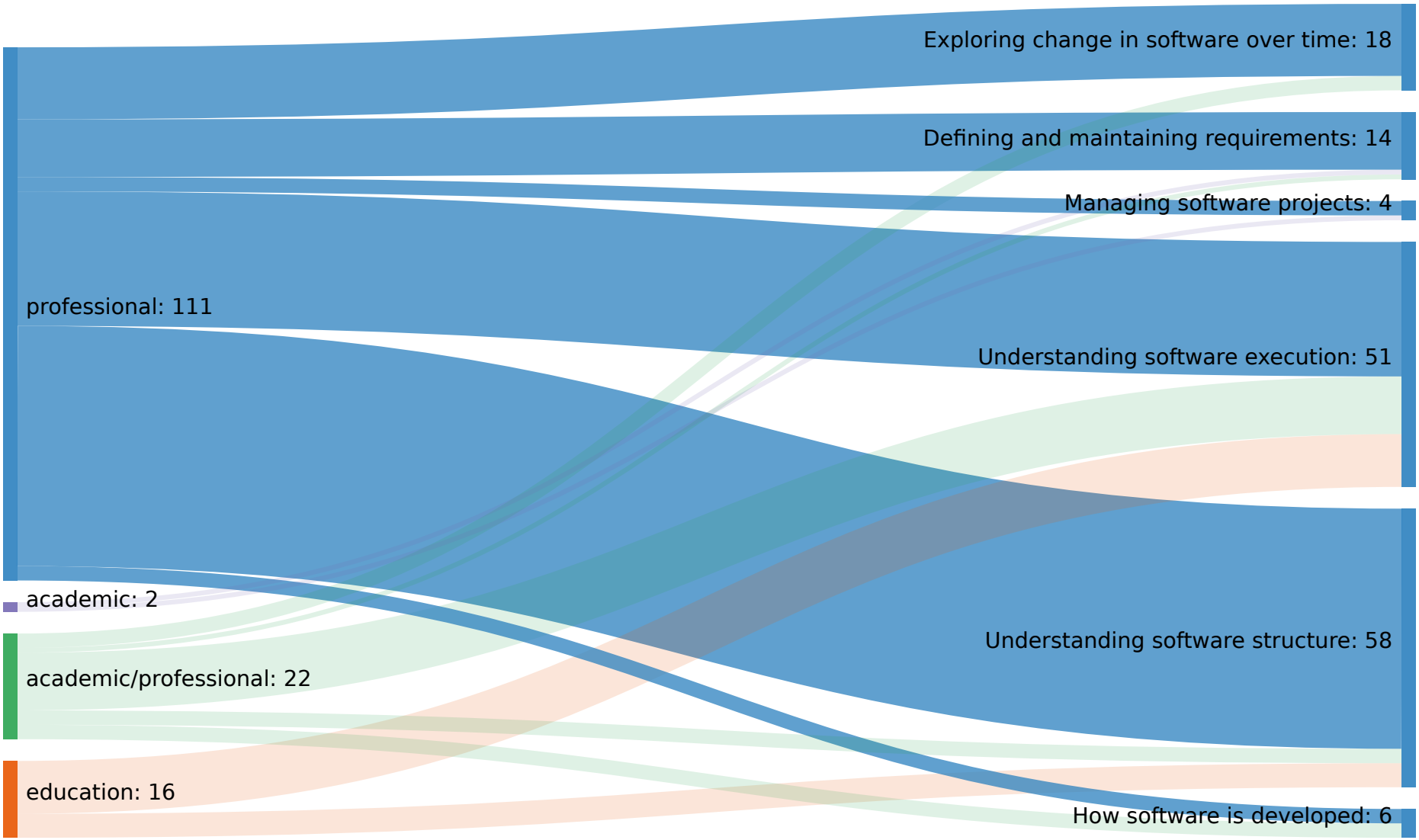
Who needs visualizations, and why?

Target audience vs. problem domain visualization contribution



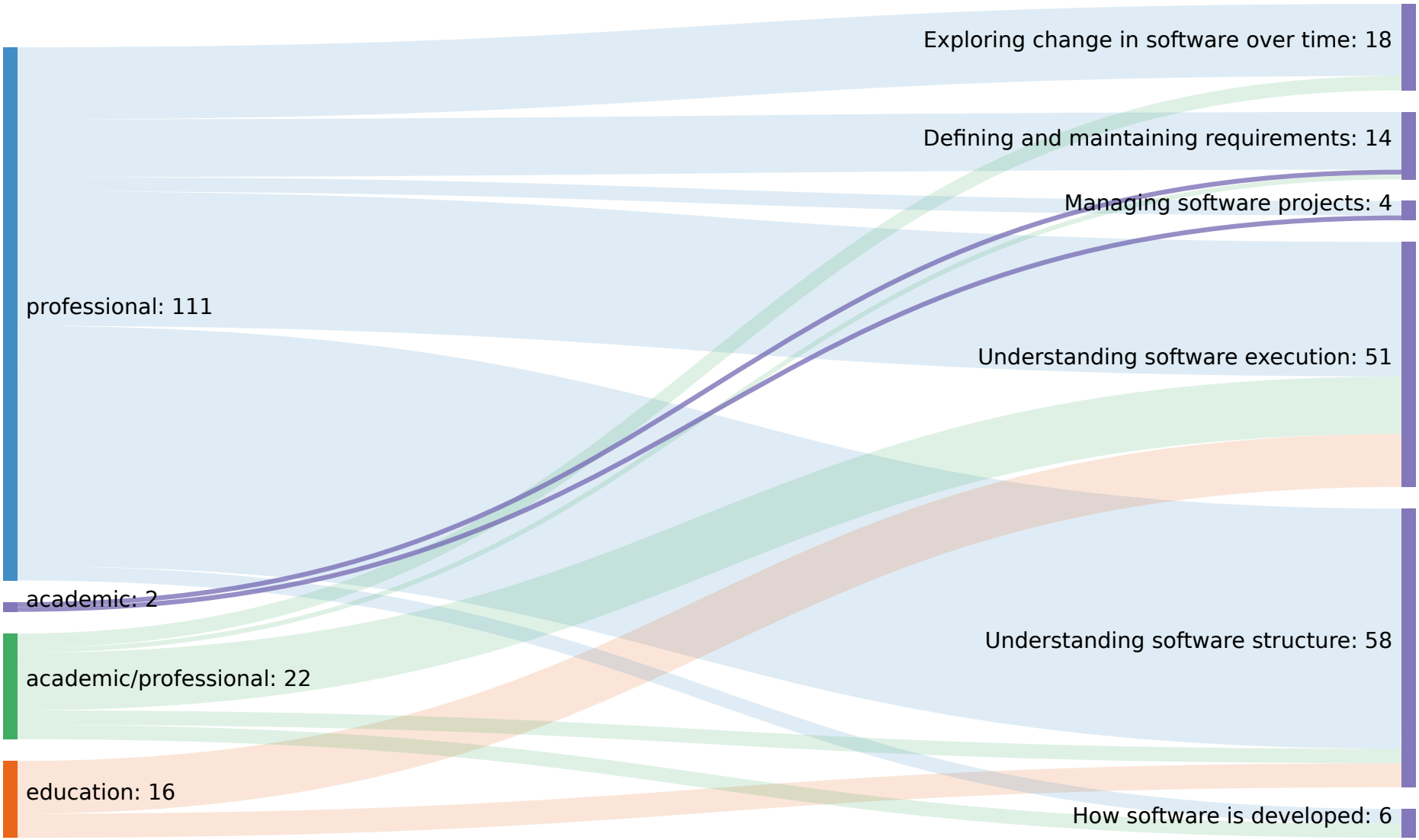
Who needs visualizations, and why?

Target audience vs. problem domain visualization contribution



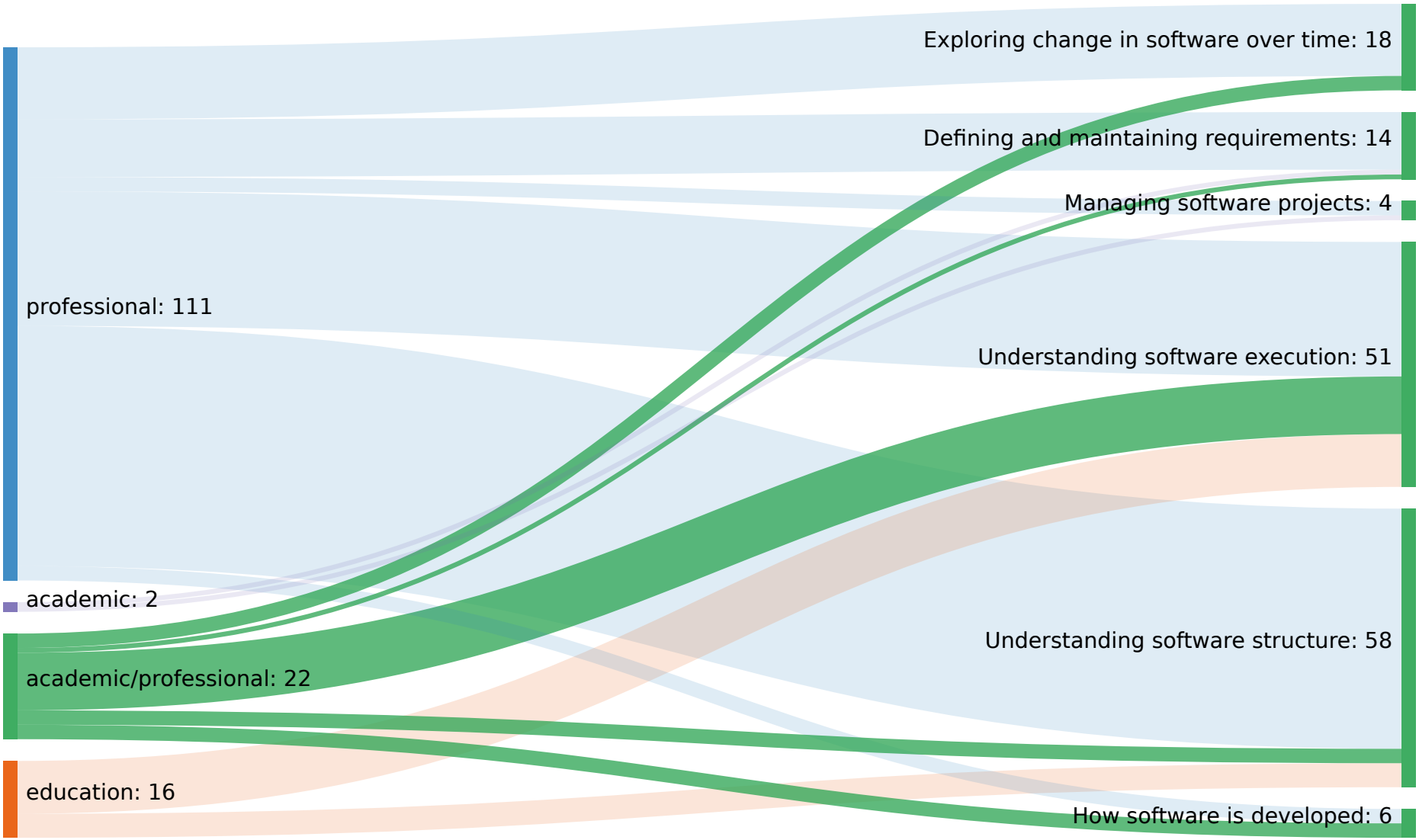
Who needs visualizations, and why?

Target audience vs. problem domain visualization contribution



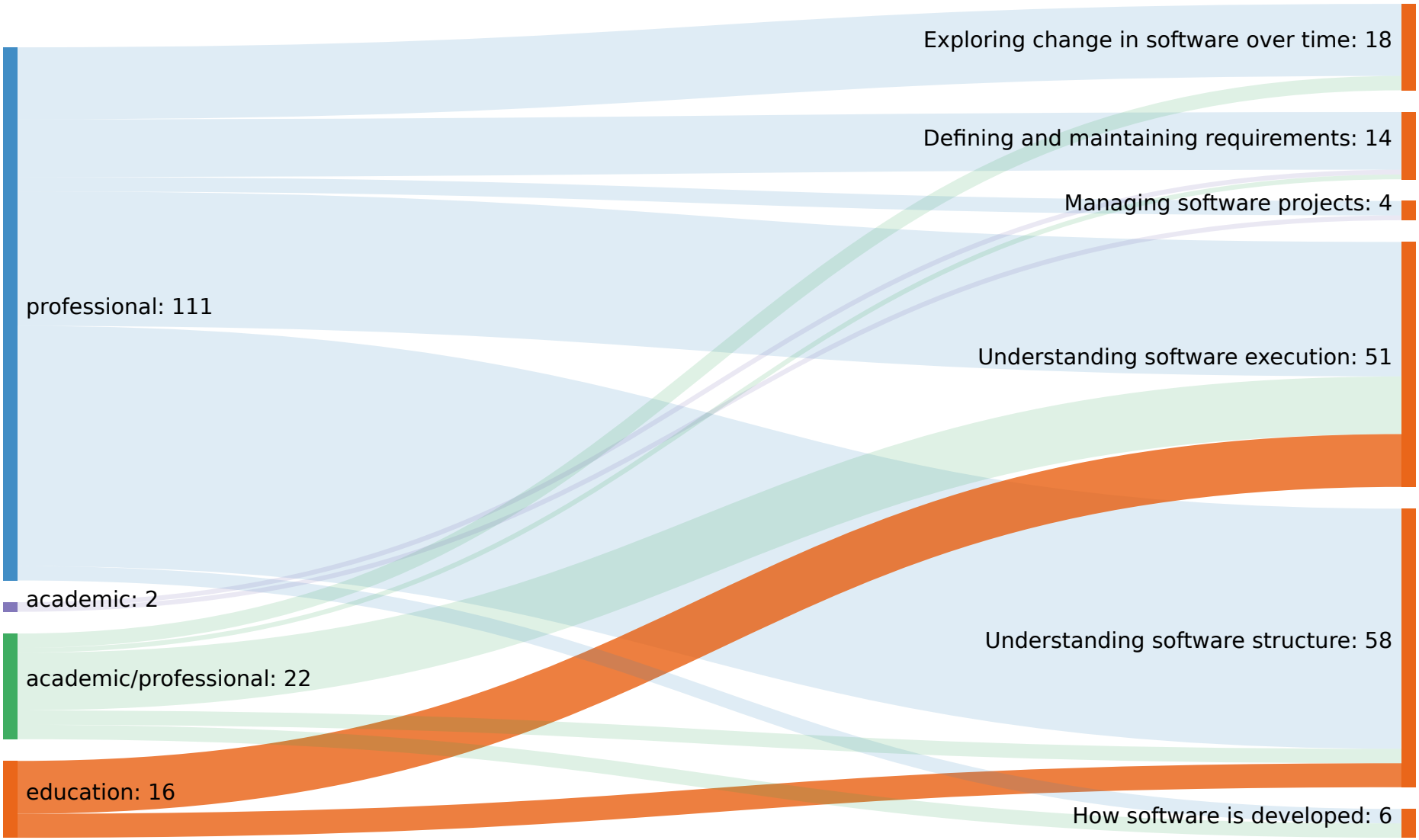
Who needs visualizations, and why?

Target audience vs. problem domain visualization contribution



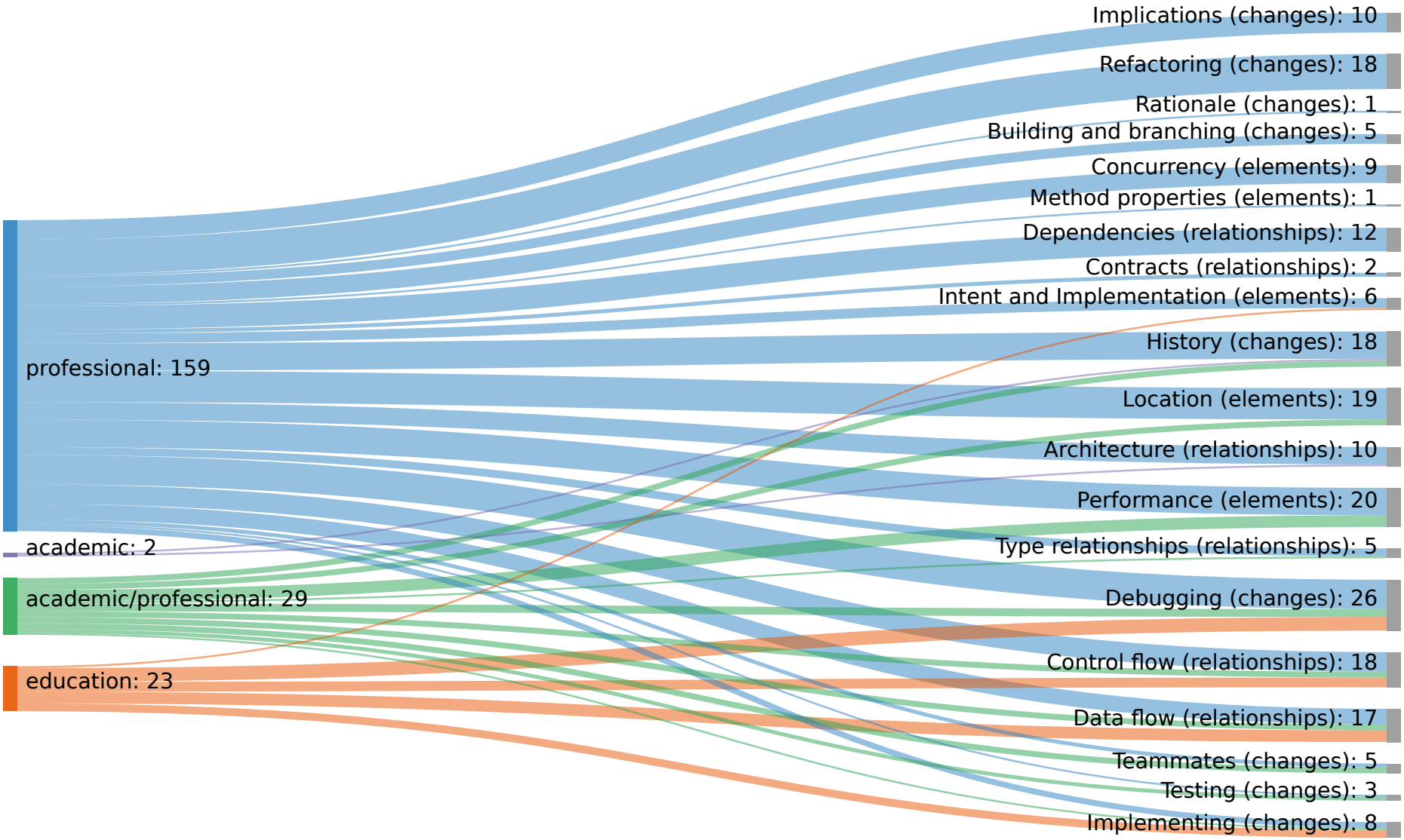
Who needs visualizations, and why?

Target audience vs. problem domain visualization contribution



Who needs visualizations, and why?

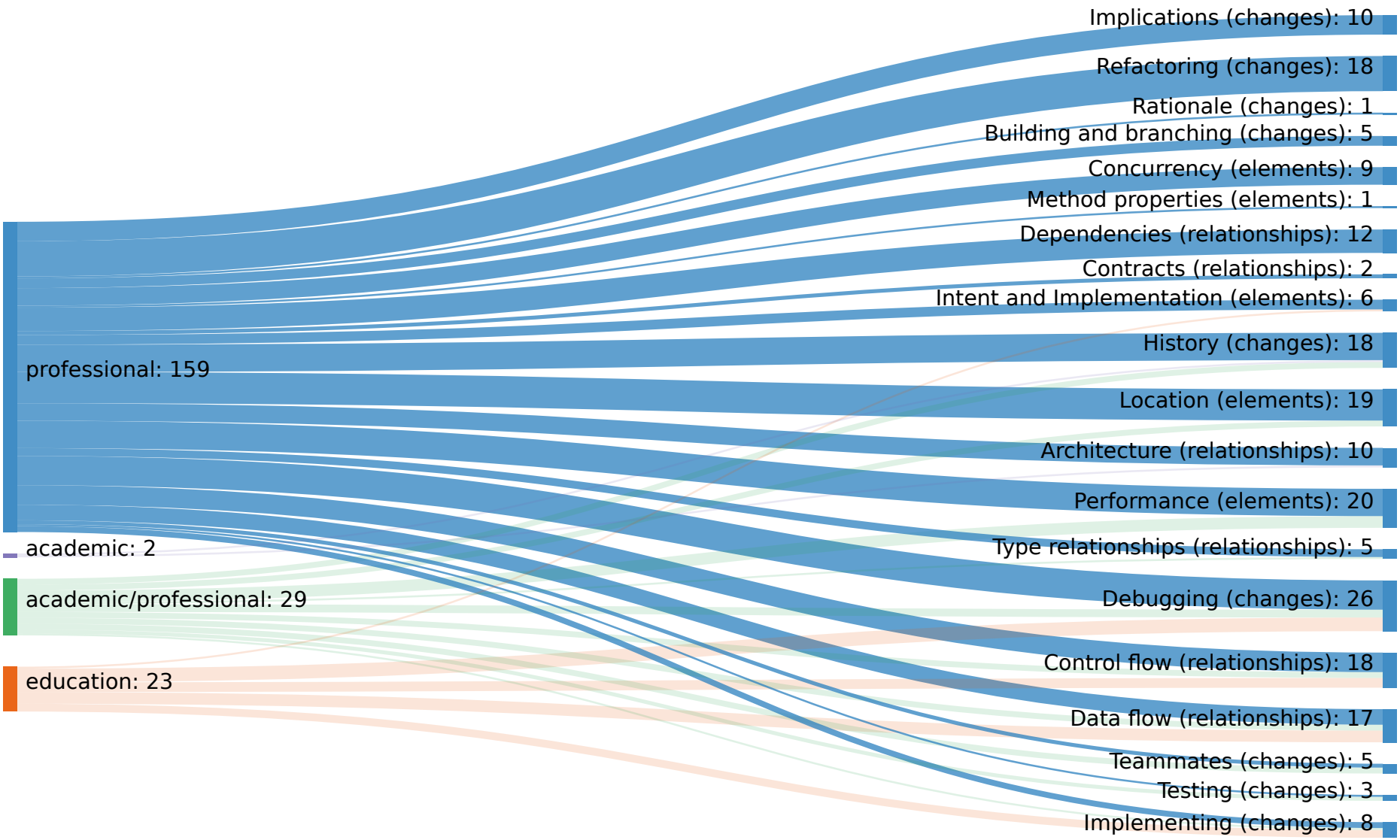
Target audience vs. data visualization questions



Data visualization questions by LaToza, Myers (2010): *"Hard-to-answer Questions About Code"*

Who needs visualizations, and why?

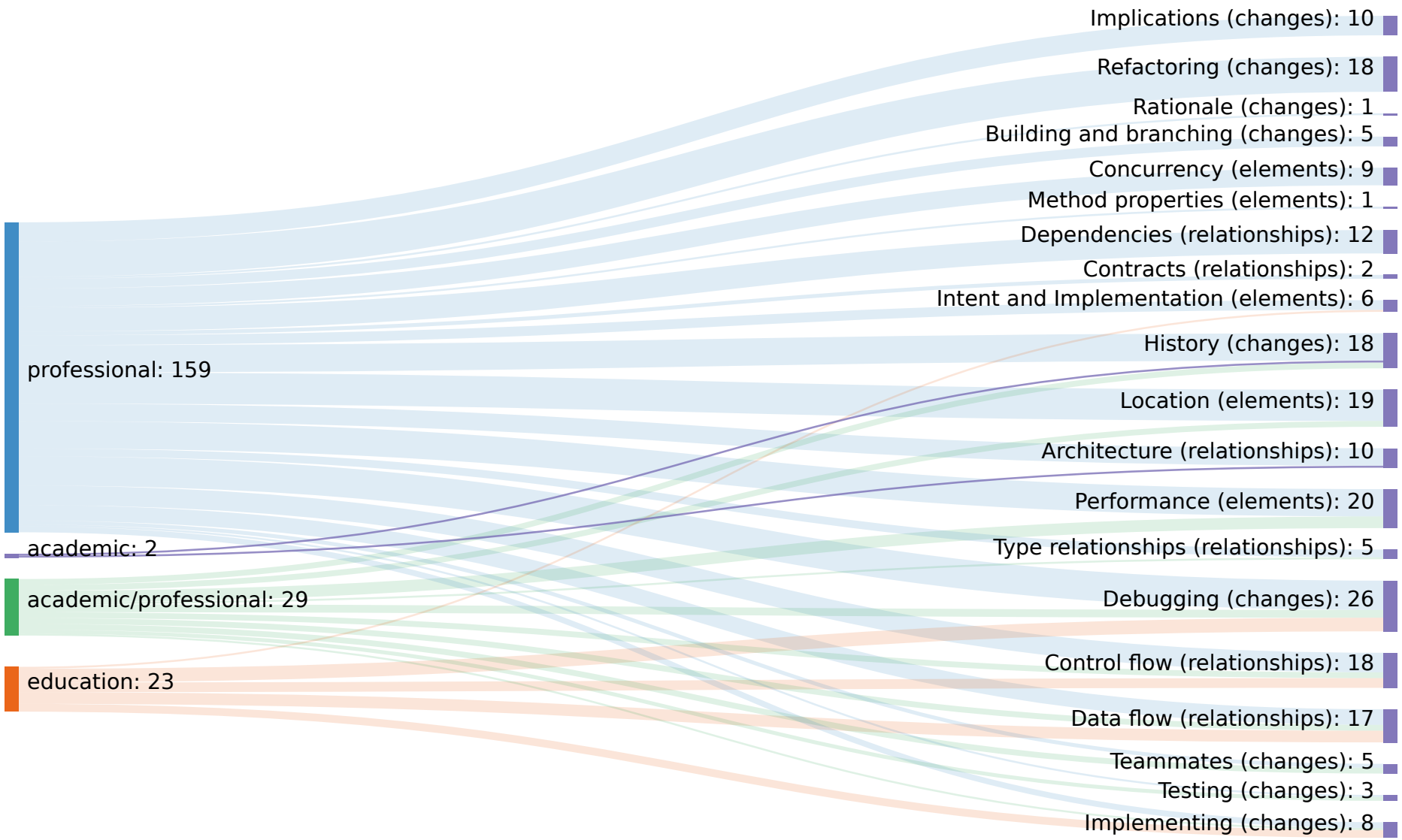
Target audience vs. data visualization questions



Data visualization questions by LaToza, Myers (2010): *"Hard-to-answer Questions About Code"*

Who needs visualizations, and why?

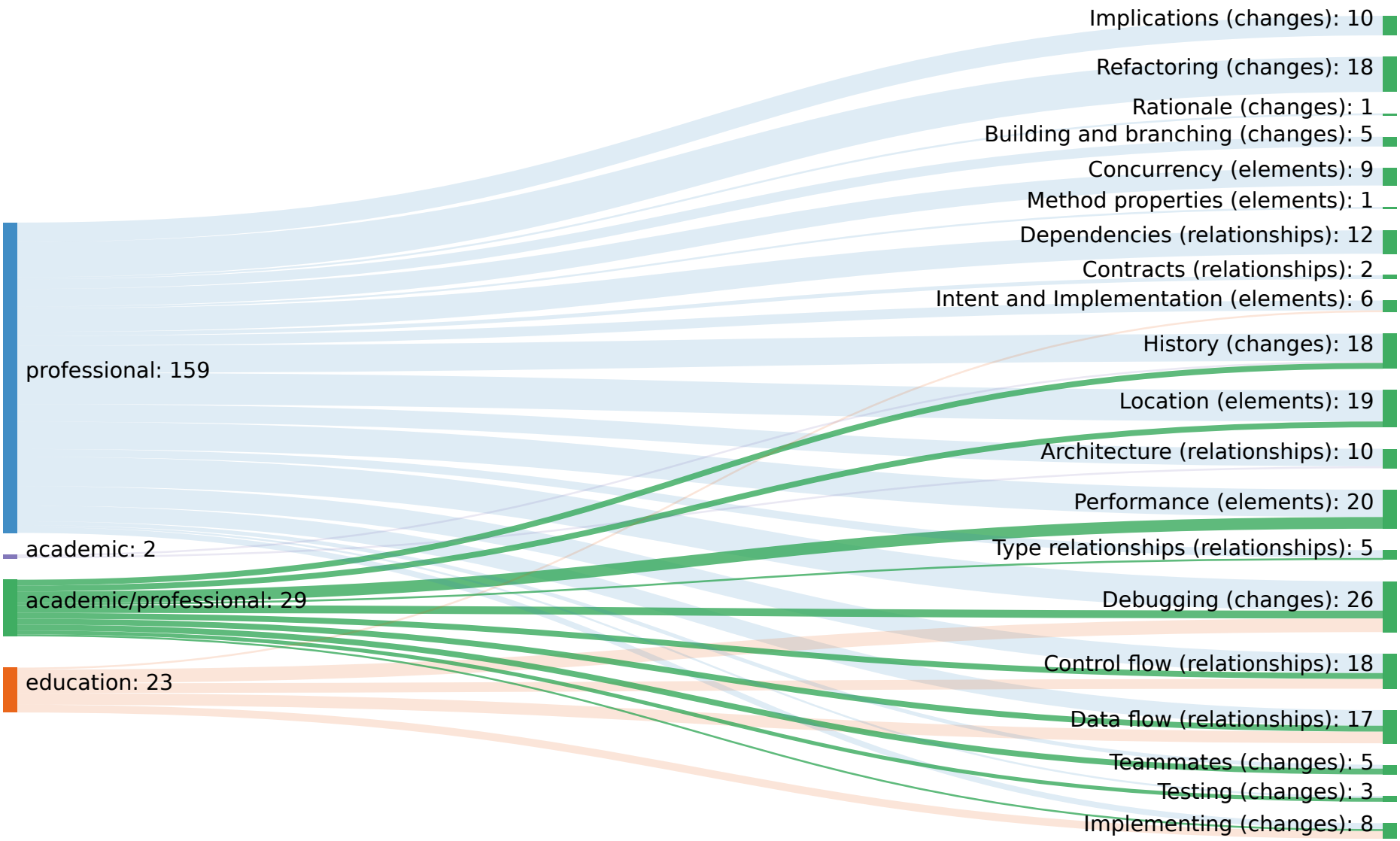
Target audience vs. data visualization questions



Data visualization questions by LaToza, Myers (2010): *"Hard-to-answer Questions About Code"*

Who needs visualizations, and why?

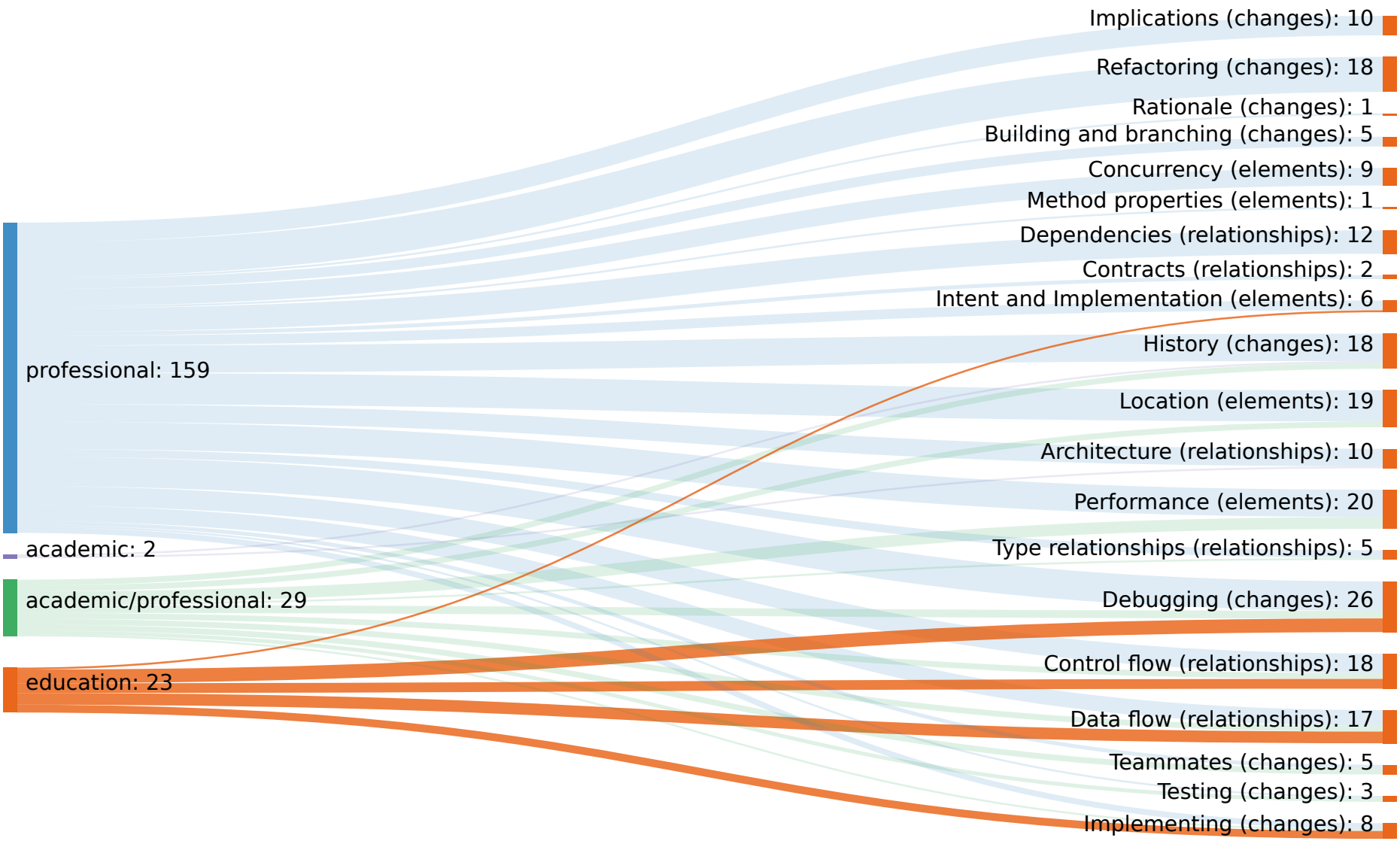
Target audience vs. data visualization questions



Data visualization questions by LaToza, Myers (2010): *"Hard-to-answer Questions About Code"*

Who needs visualizations, and why?

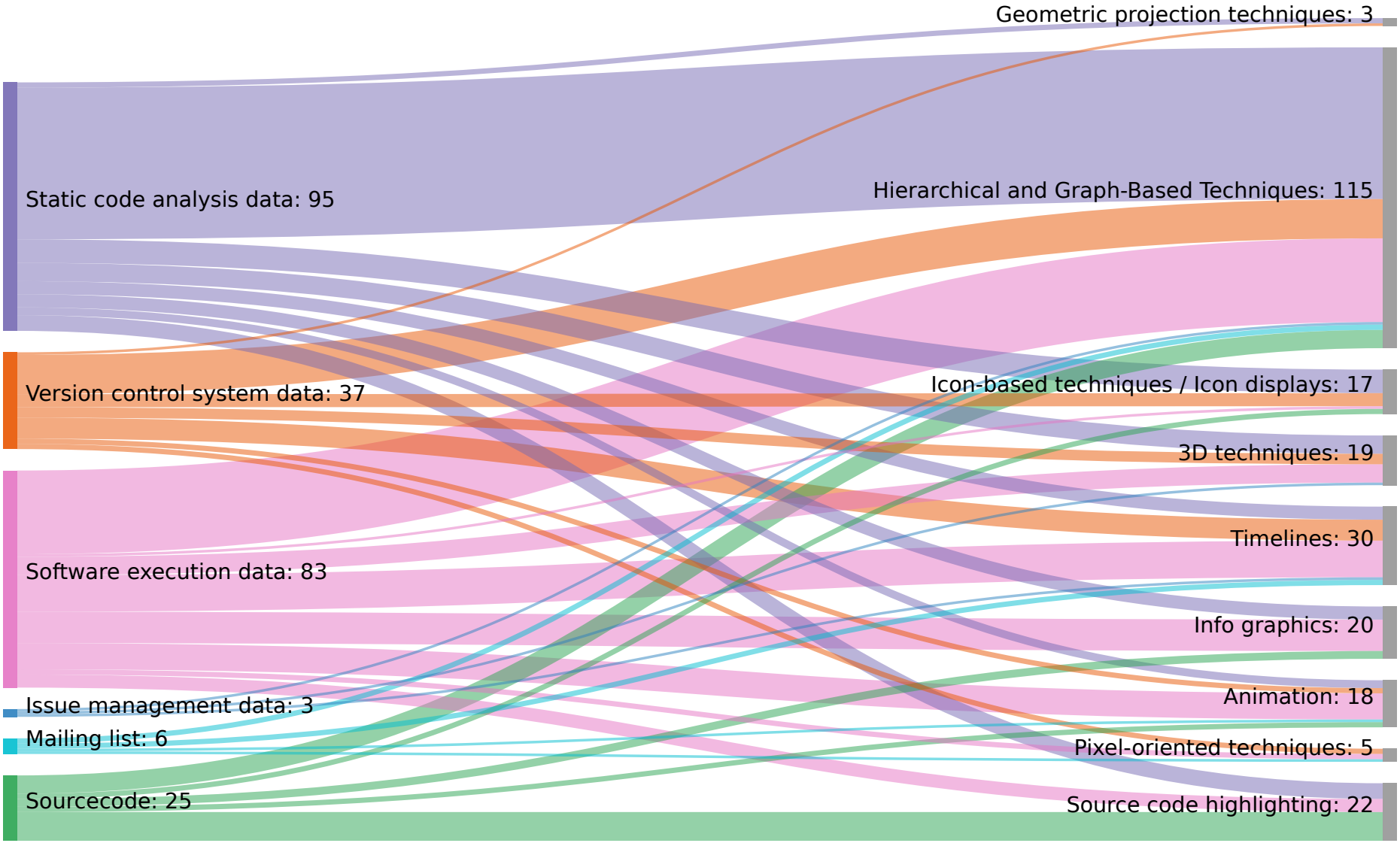
Target audience vs. data visualization questions



Data visualization questions by LaToza, Myers (2010): *"Hard-to-answer Questions About Code"*

What are the sources for which visualizations?

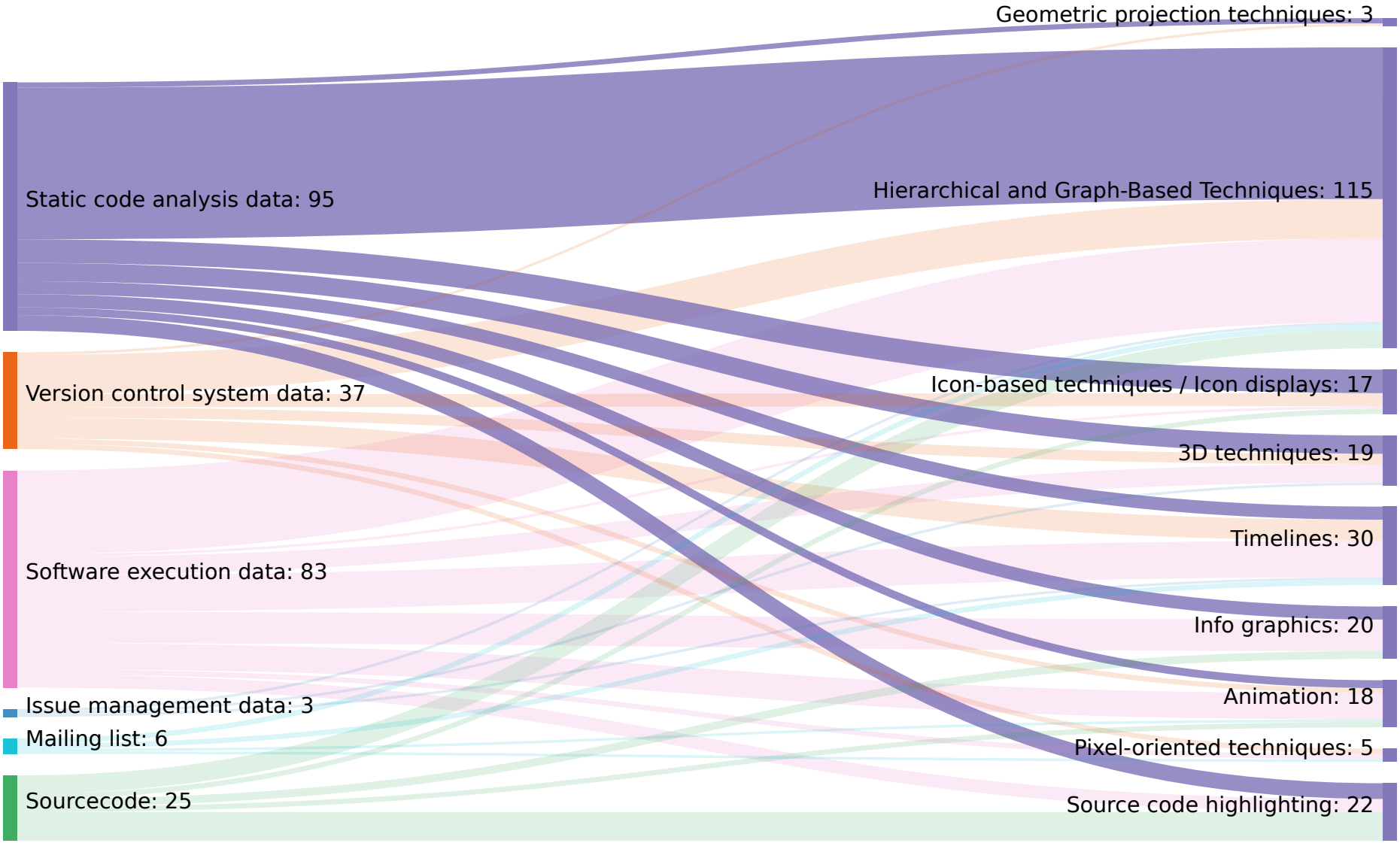
Visualzation sources vs. visualization paradigms



Visualization Paradigms by Keim, Kriegel (1996): *"Visualization Techniques for Mining Large Databases: A Comparison"*

What are the sources for which visualizations?

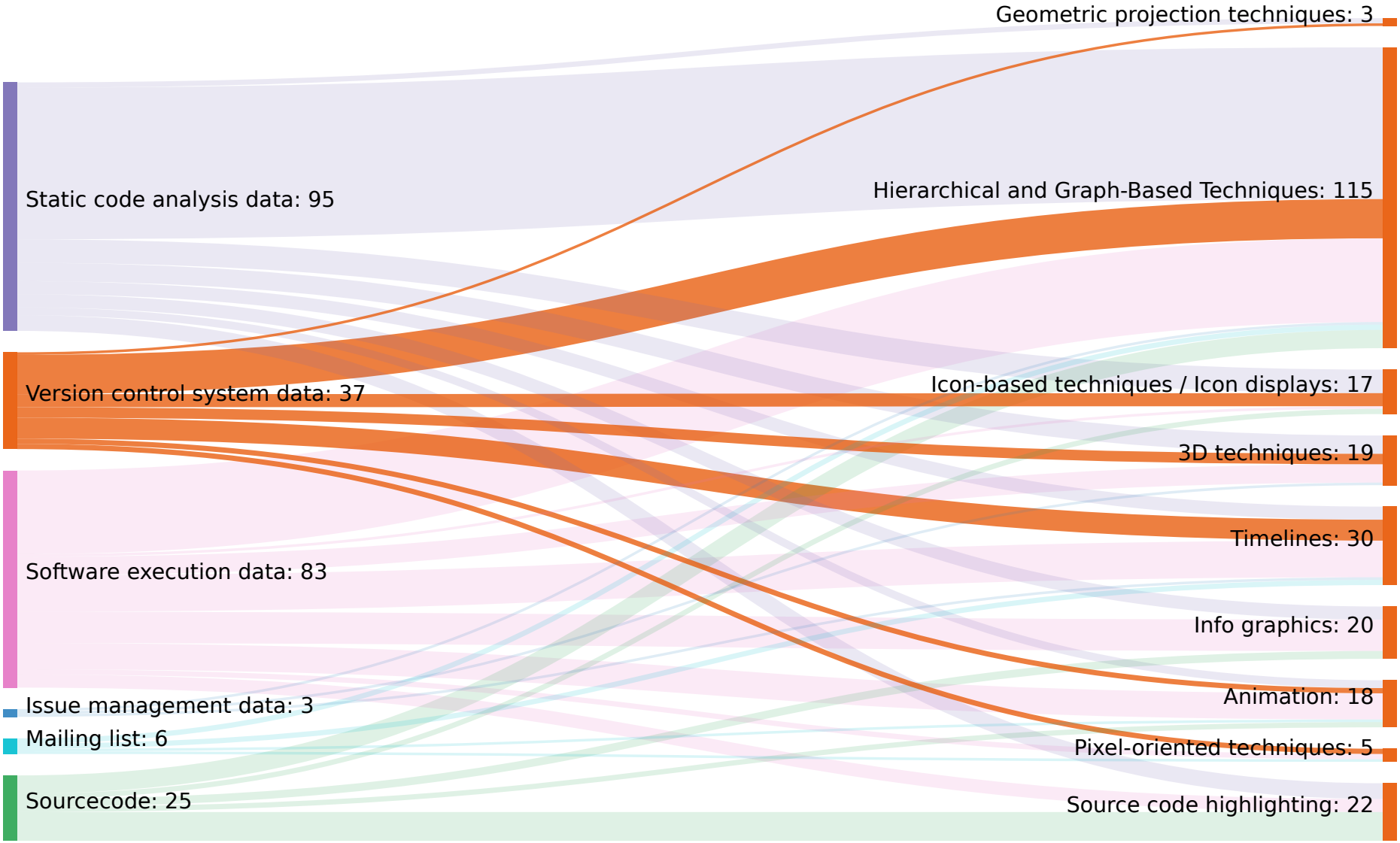
Visualzation sources vs. visualization paradigms



Visualization Paradigms by Keim, Kriegel (1996): *"Visualization Techniques for Mining Large Databases: A Comparison"*

What are the sources for which visualizations?

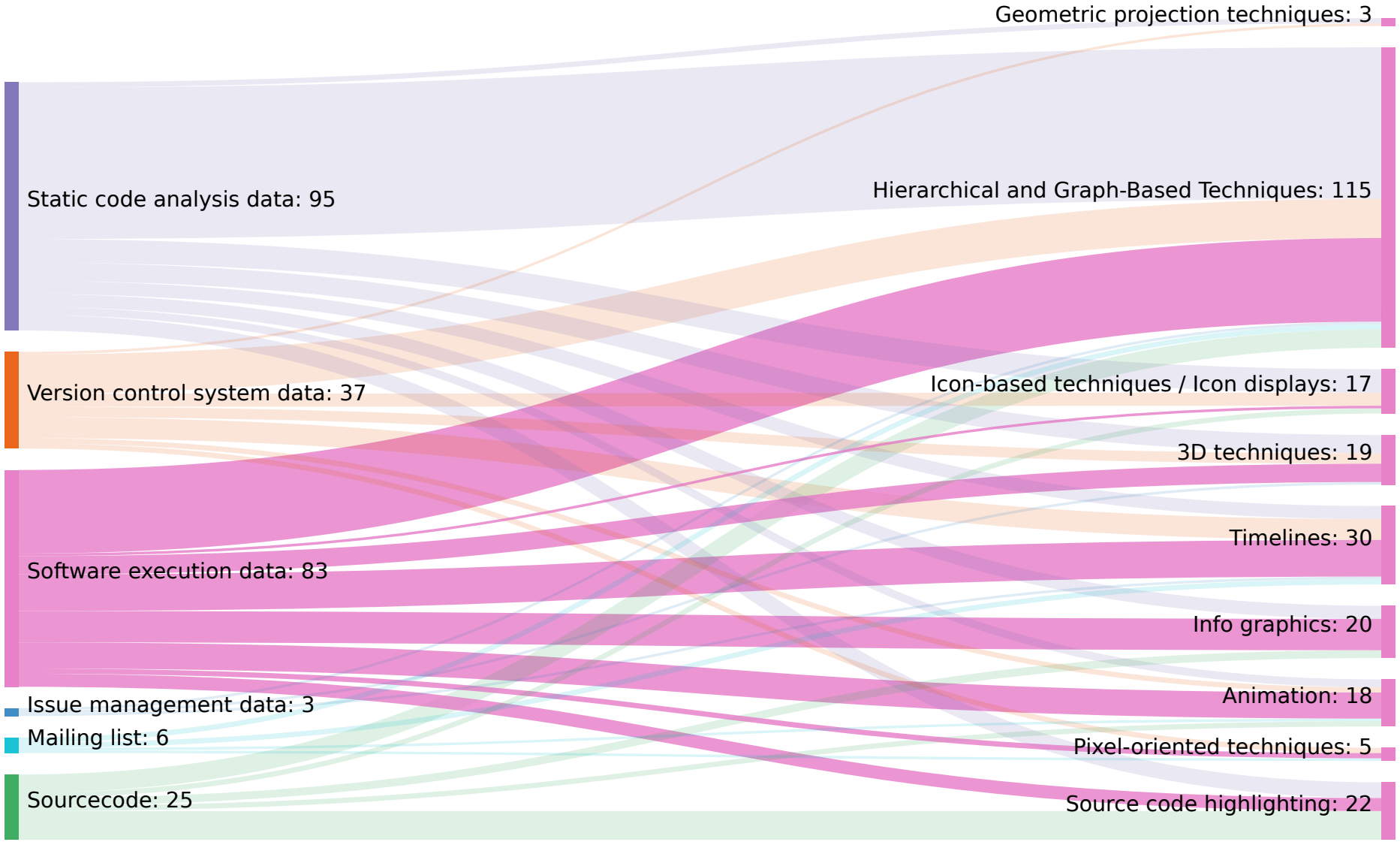
Visualzation sources vs. visualization paradigms



Visualization Paradigms by Keim, Kriegel (1996): *"Visualization Techniques for Mining Large Databases: A Comparison"*

What are the sources for which visualizations?

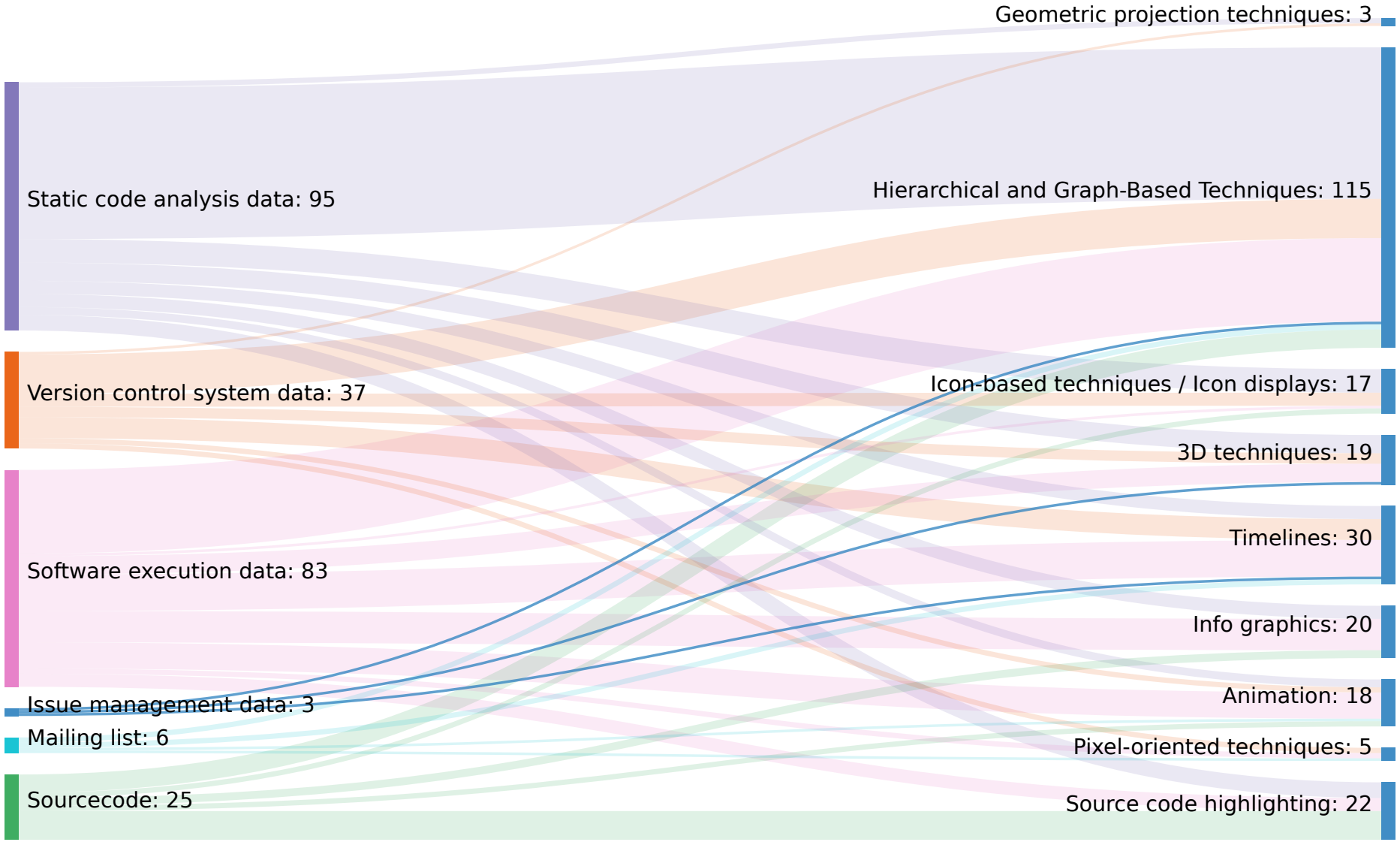
Visualzation sources vs. visualization paradigms



Visualization Paradigms by Keim, Kriegel (1996): *"Visualization Techniques for Mining Large Databases: A Comparison"*

What are the sources for which visualizations?

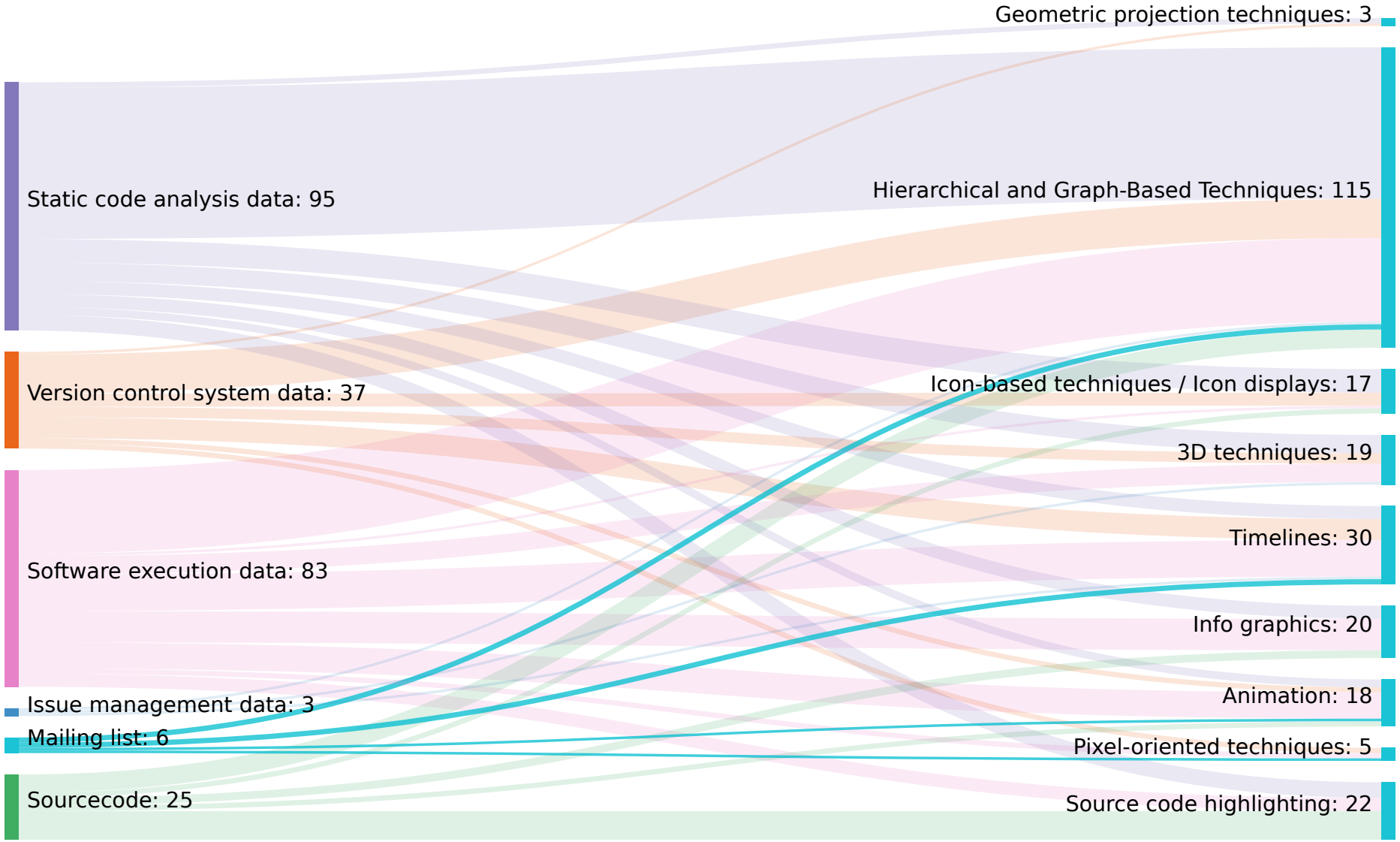
Visualzation sources vs. visualization paradigms



Visualization Paradigms by Keim, Kriegel (1996): *"Visualization Techniques for Mining Large Databases: A Comparison"*

What are the sources for which visualizations?

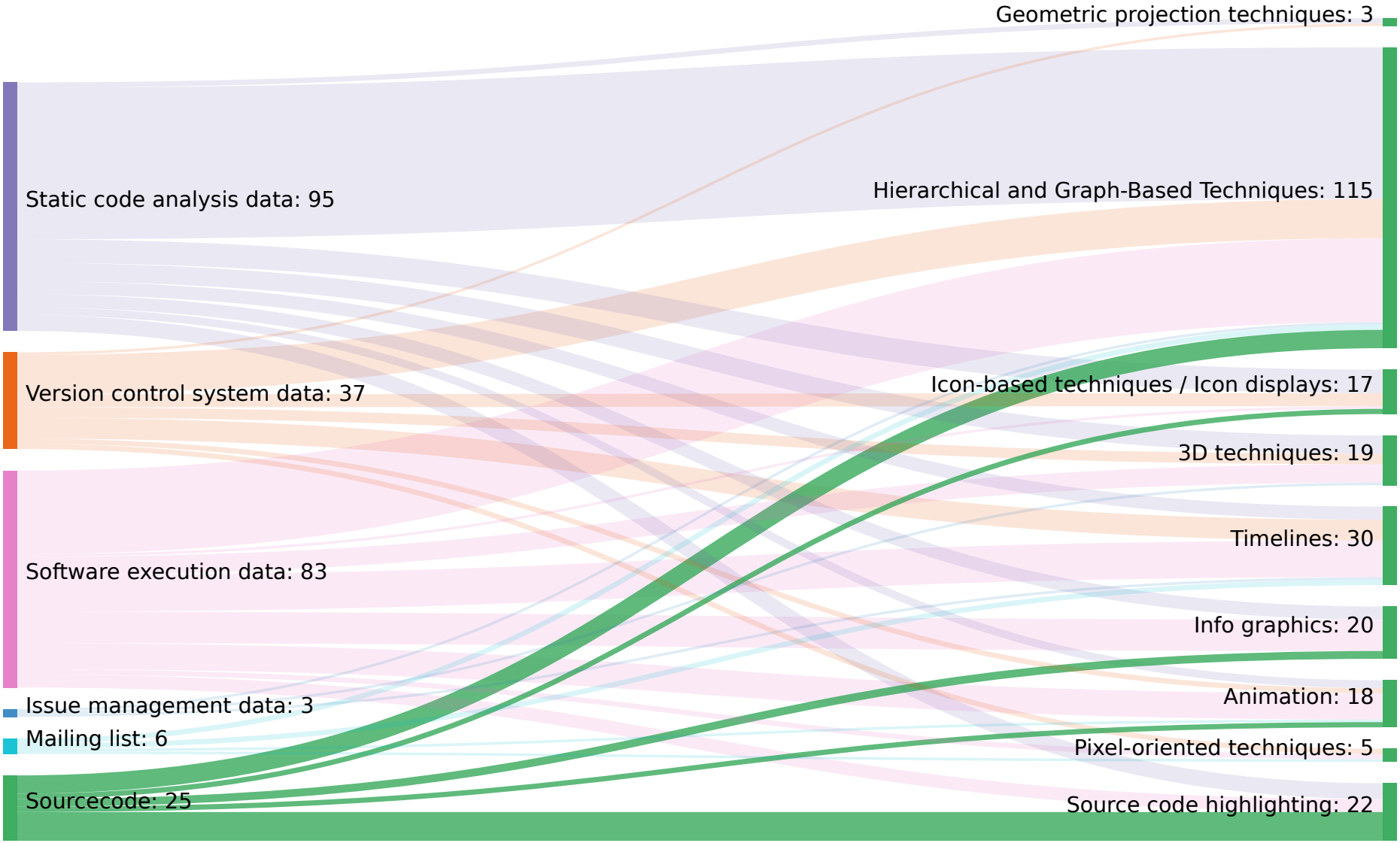
Visualzation sources vs. visualization paradigms



Visualization Paradigms by Keim, Kriegel (1996): *"Visualization Techniques for Mining Large Databases: A Comparison"*

What are the sources for which visualizations?

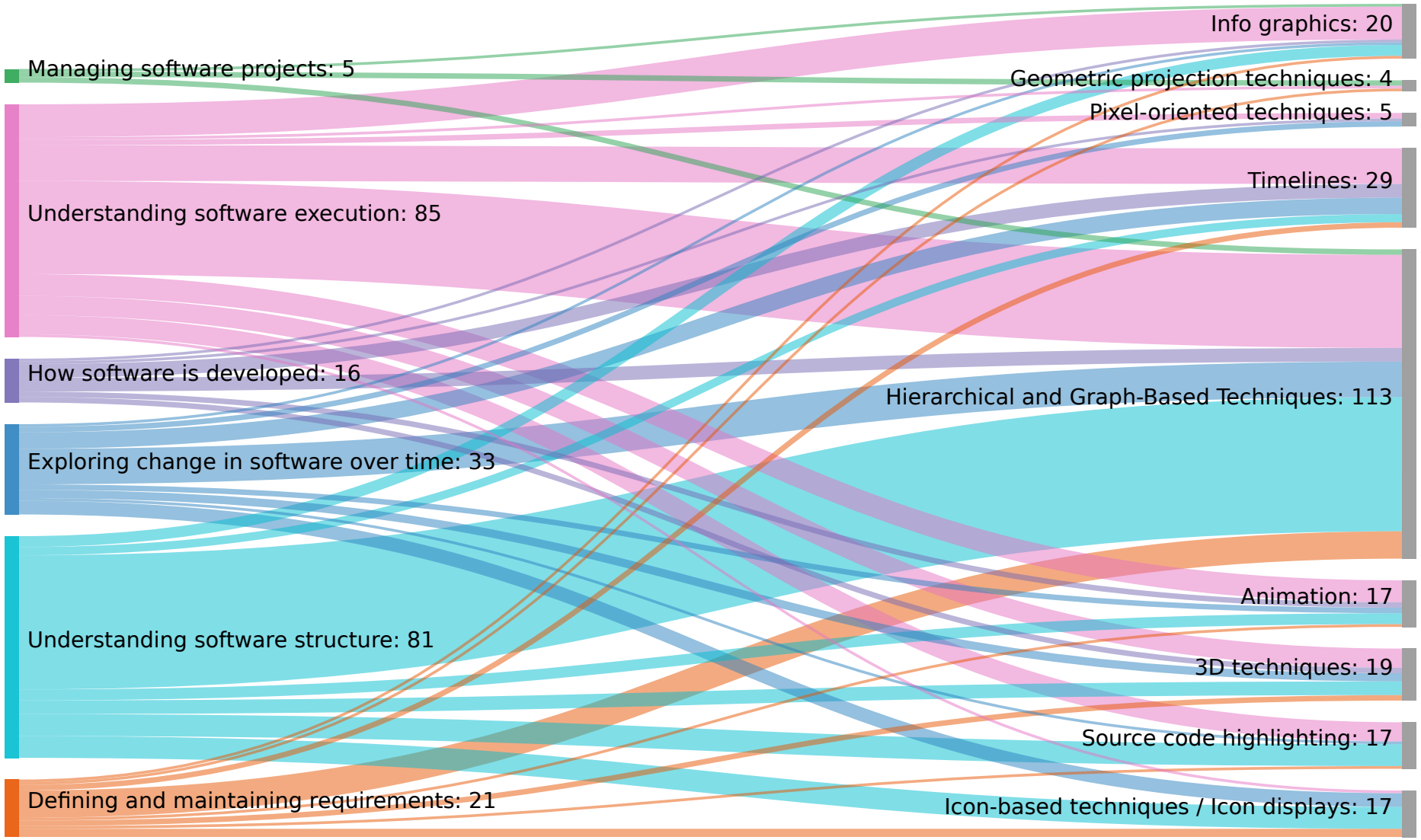
Visualzation sources vs. visualization paradigms



Visualization Paradigms by Keim, Kriegel (1996): *"Visualization Techniques for Mining Large Databases: A Comparison"*

What is visualized?

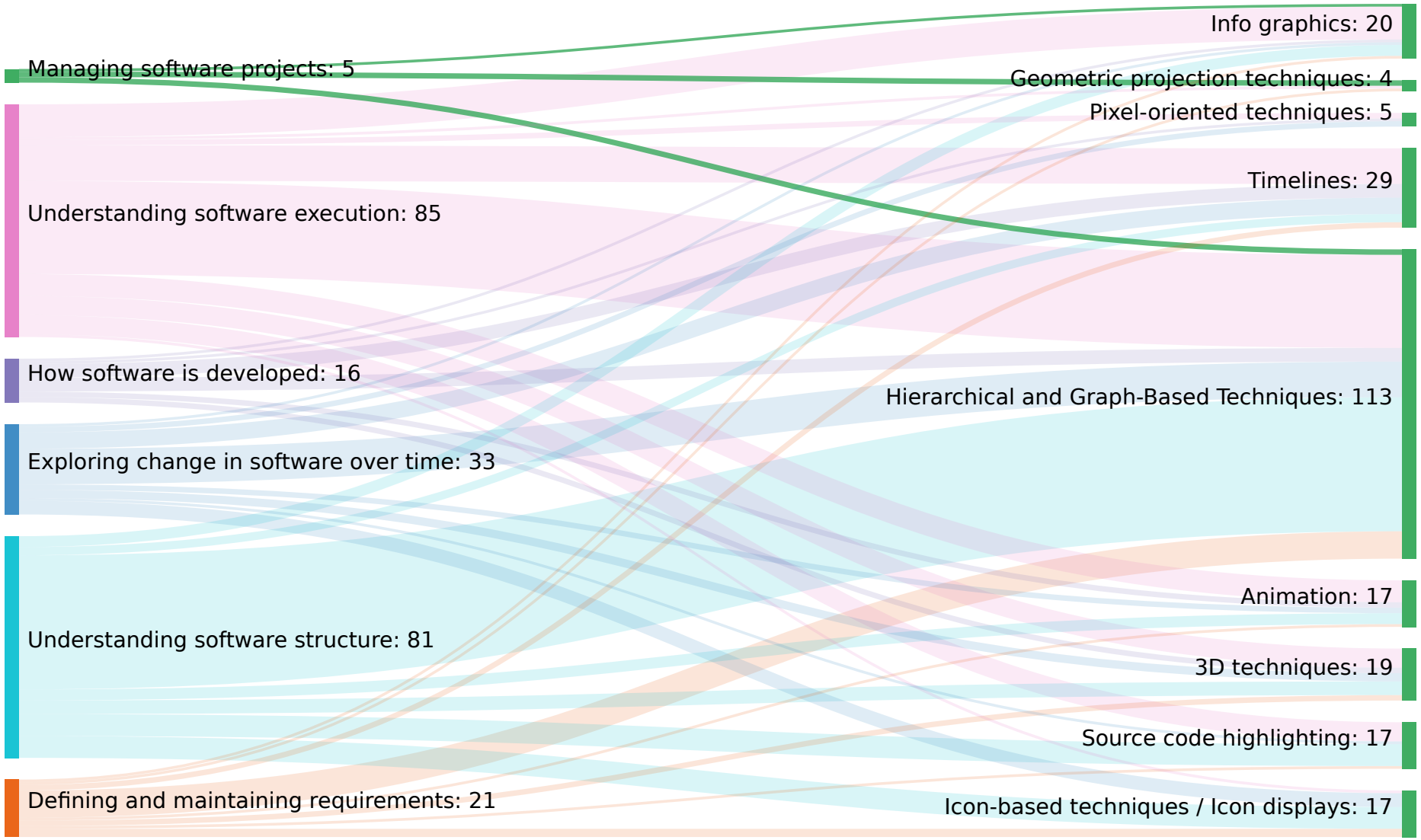
Problem domain visualization contribution vs. visualization paradigms



Visualization Paradigms by Keim, Kriegel (1996): *"Visualization Techniques for Mining Large Databases: A Comparison"*

What is visualized?

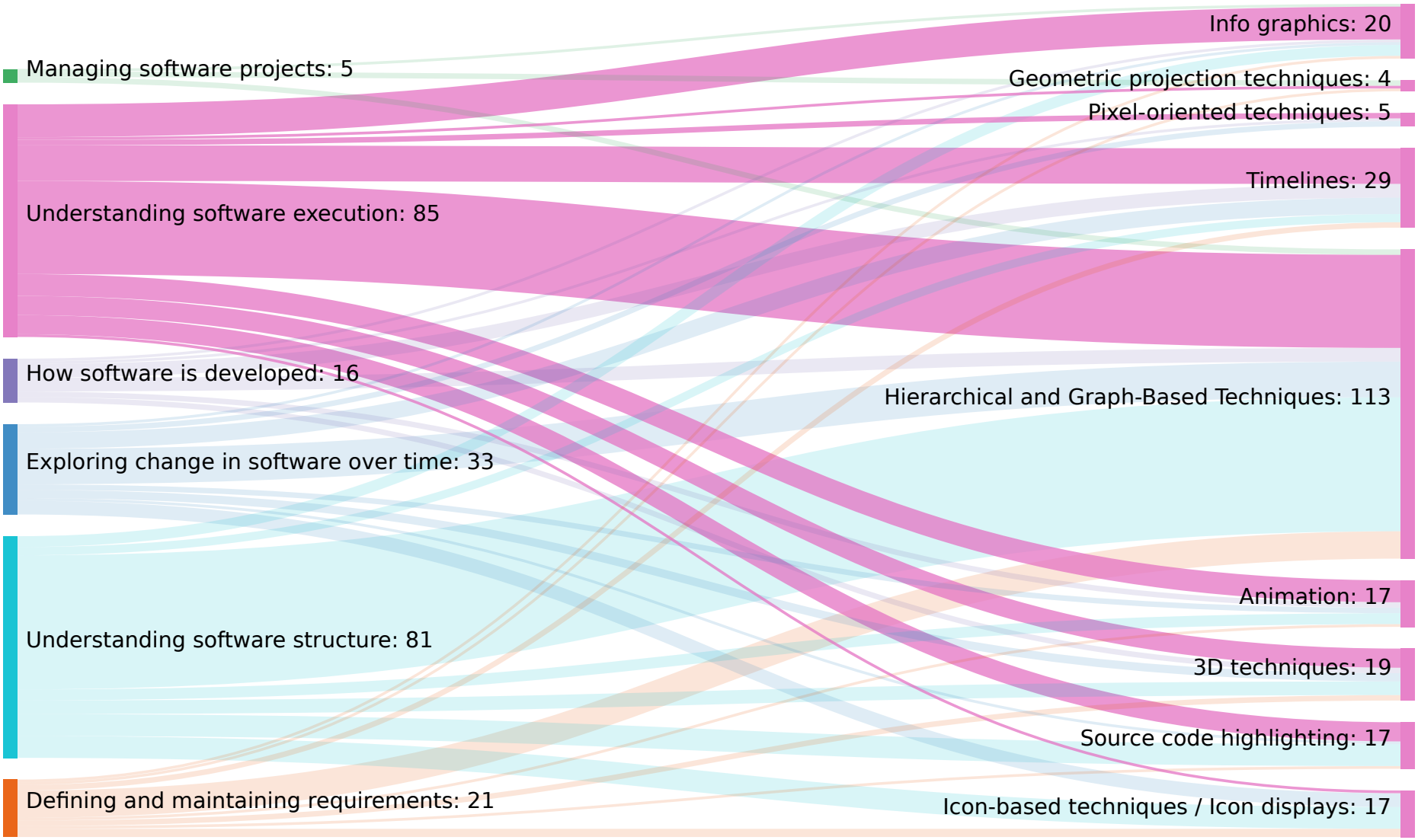
Problem domain visualization contribution vs. visualization paradigms



Visualization Paradigms by Keim, Kriegel (1996): *"Visualization Techniques for Mining Large Databases: A Comparison"*

What is visualized?

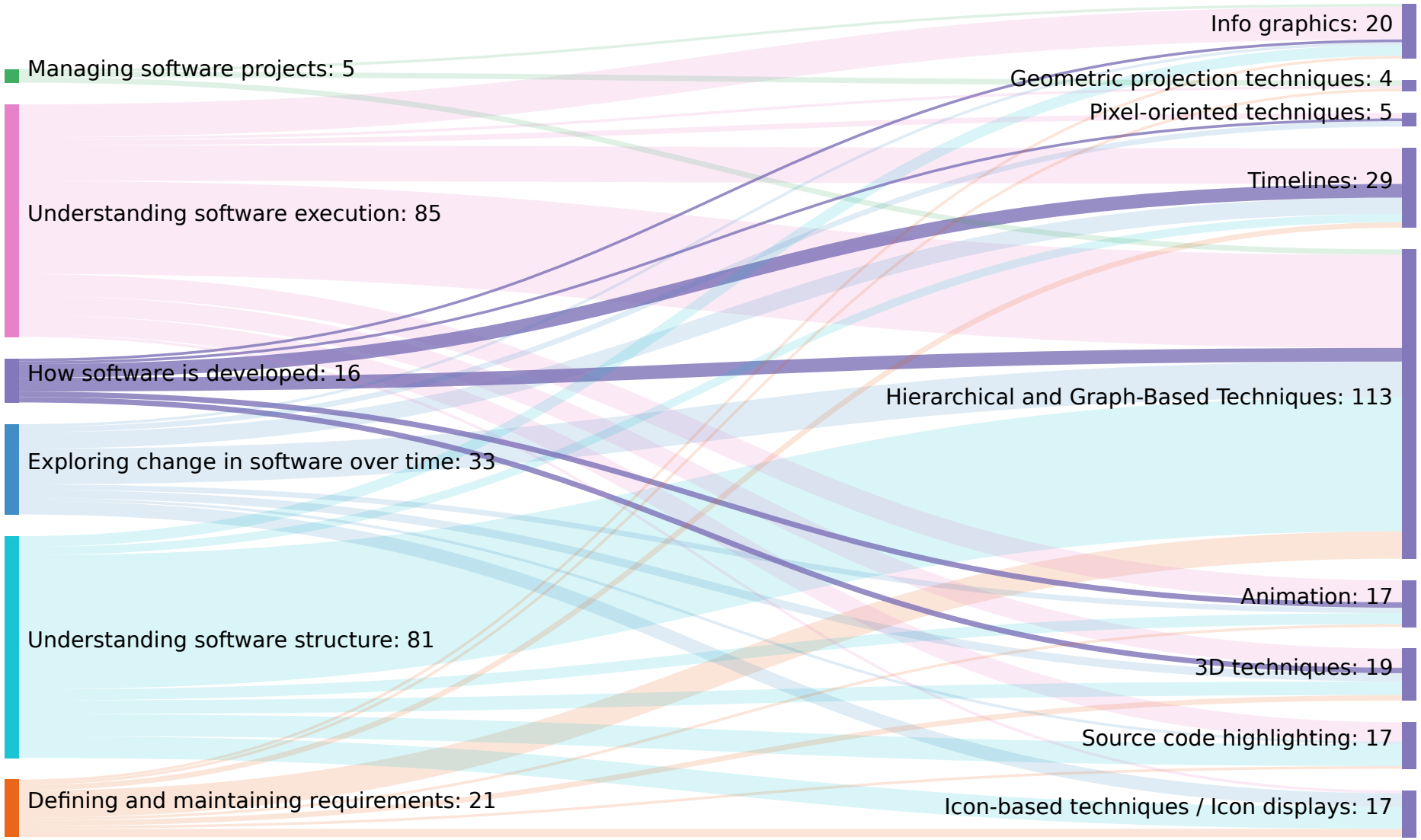
Problem domain visualization contribution vs. visualization paradigms



Visualization Paradigms by Keim, Kriegel (1996): *"Visualization Techniques for Mining Large Databases: A Comparison"*

What is visualized?

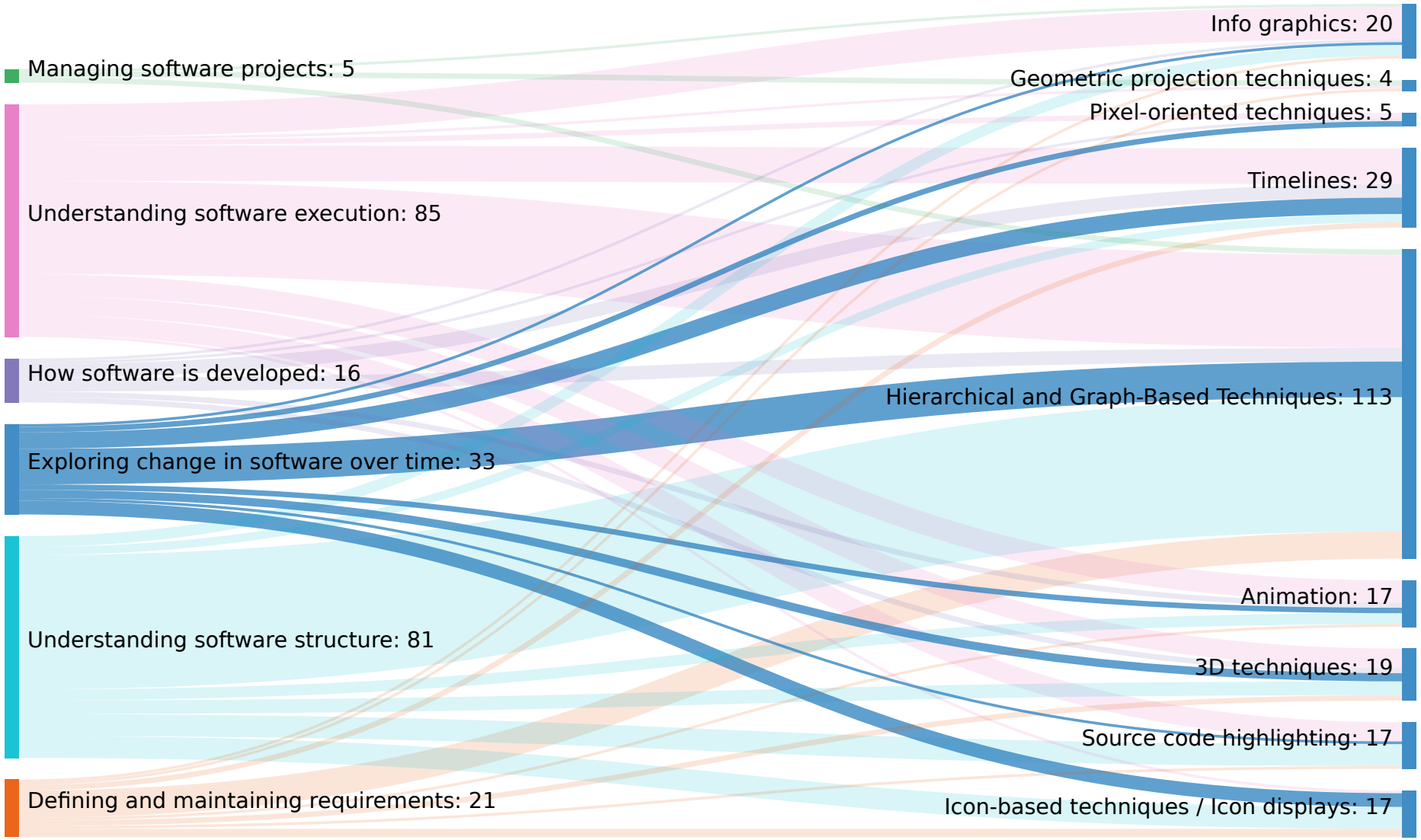
Problem domain visualization contribution vs. visualization paradigms



Visualization Paradigms by Keim, Kriegel (1996): *"Visualization Techniques for Mining Large Databases: A Comparison"*

What is visualized?

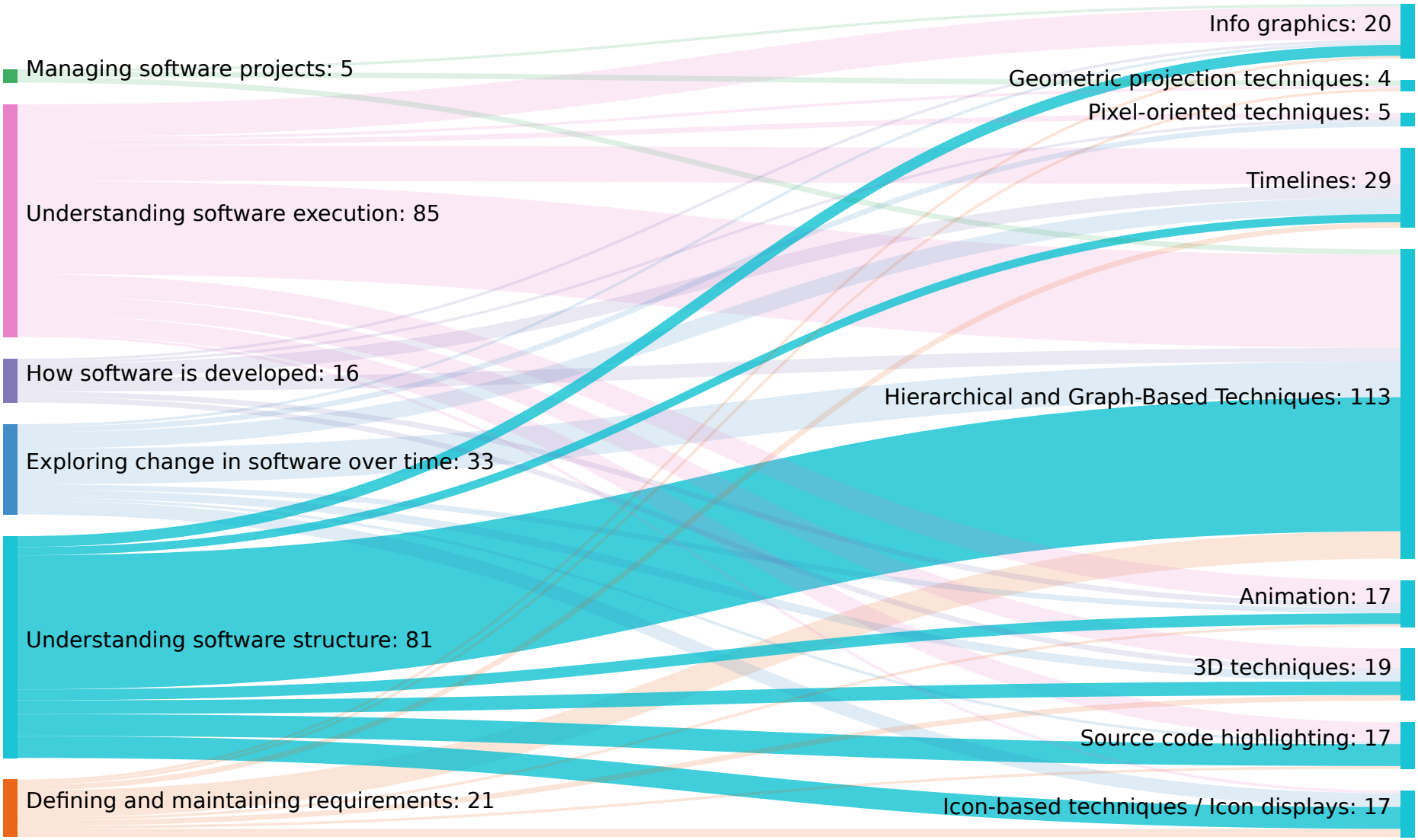
Problem domain visualization contribution vs. visualization paradigms



Visualization Paradigms by Keim, Kriegel (1996): *"Visualization Techniques for Mining Large Databases: A Comparison"*

What is visualized?

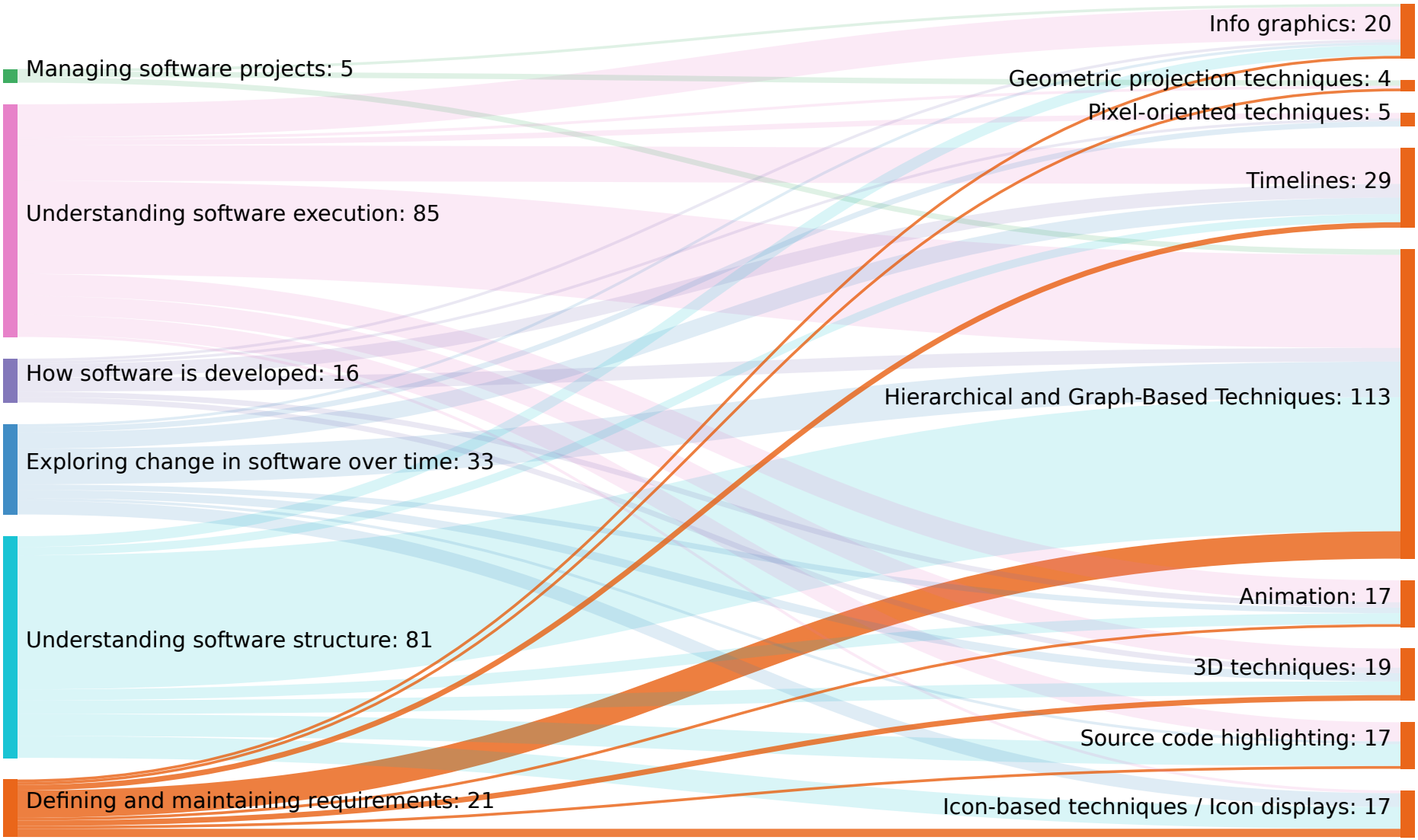
Problem domain visualization contribution vs. visualization paradigms



Visualization Paradigms by Keim, Kriegel (1996): *"Visualization Techniques for Mining Large Databases: A Comparison"*

What is visualized?

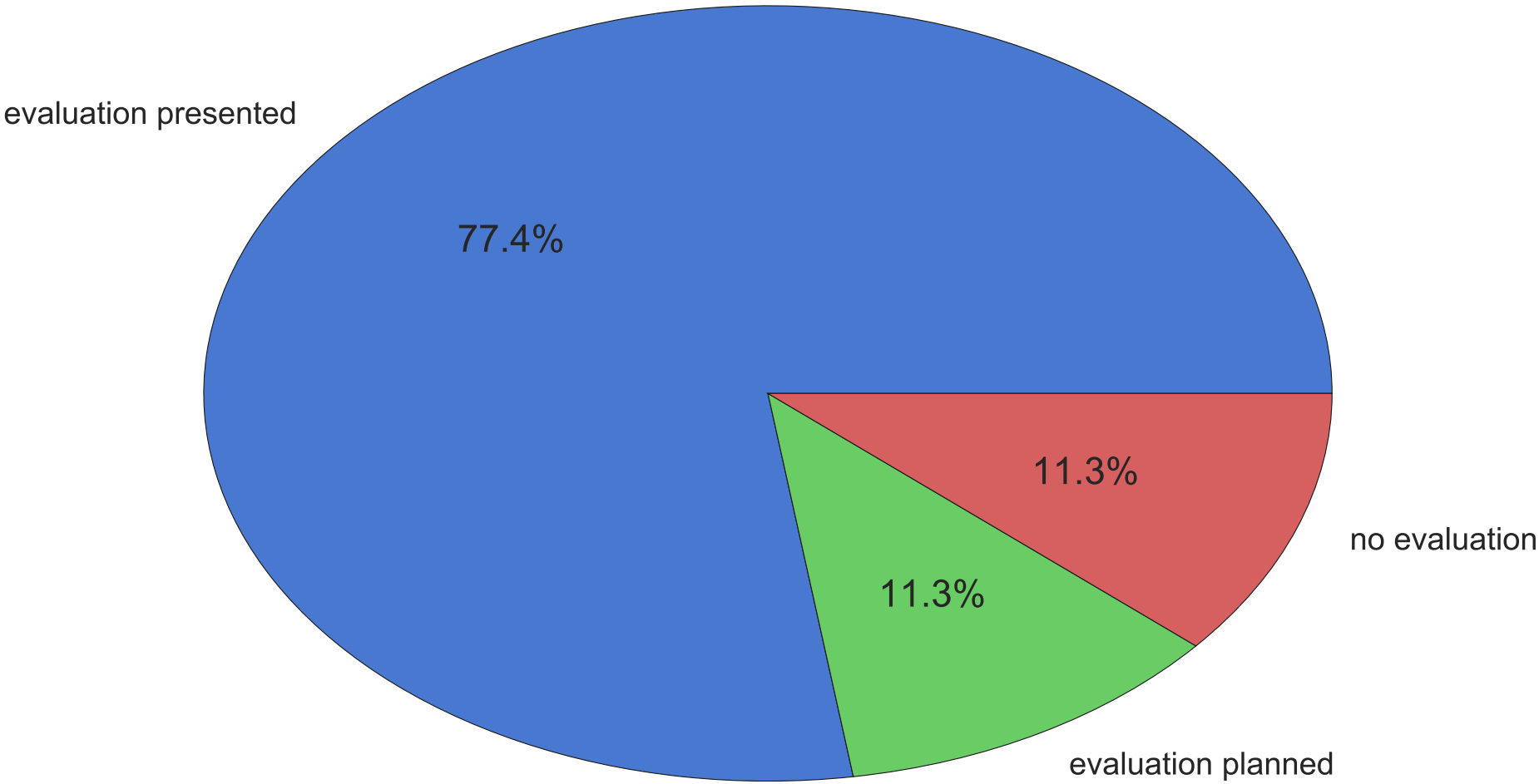
Problem domain visualization contribution vs. visualization paradigms



Visualization Paradigms by Keim, Kriegel (1996): *"Visualization Techniques for Mining Large Databases: A Comparison"*

79 Evaluations of 62 design studies

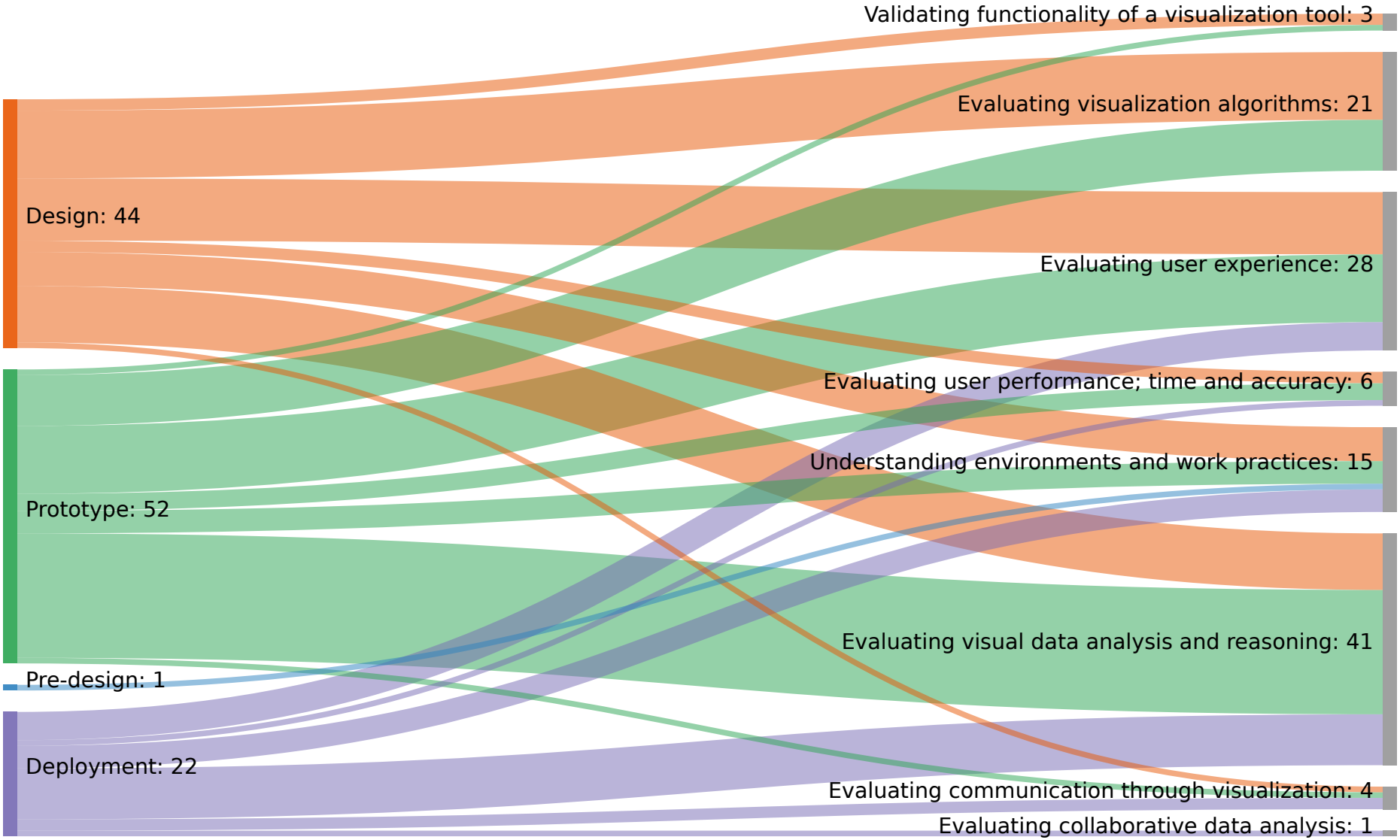
To evaluate, or not to evaluate. . .



percentage of selected design studies ($N = 62$)

What is evaluated, and when?

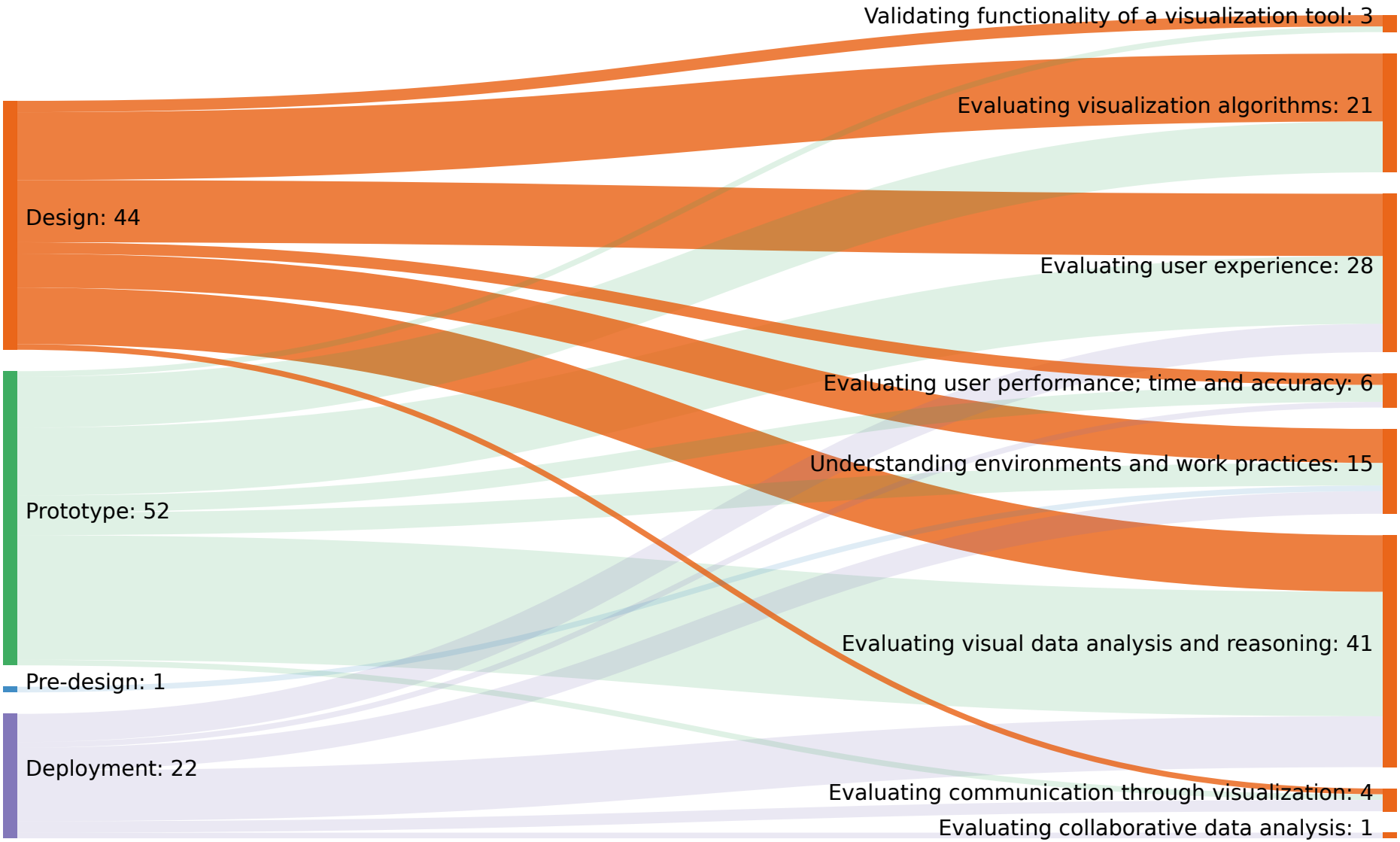
Evaluation scope vs. evaluation aspects



Evaluation aspects by Lam et al. (2012): *"Empirical Studies in Information Visualization: Seven Scenarios"*

What is evaluated, and when?

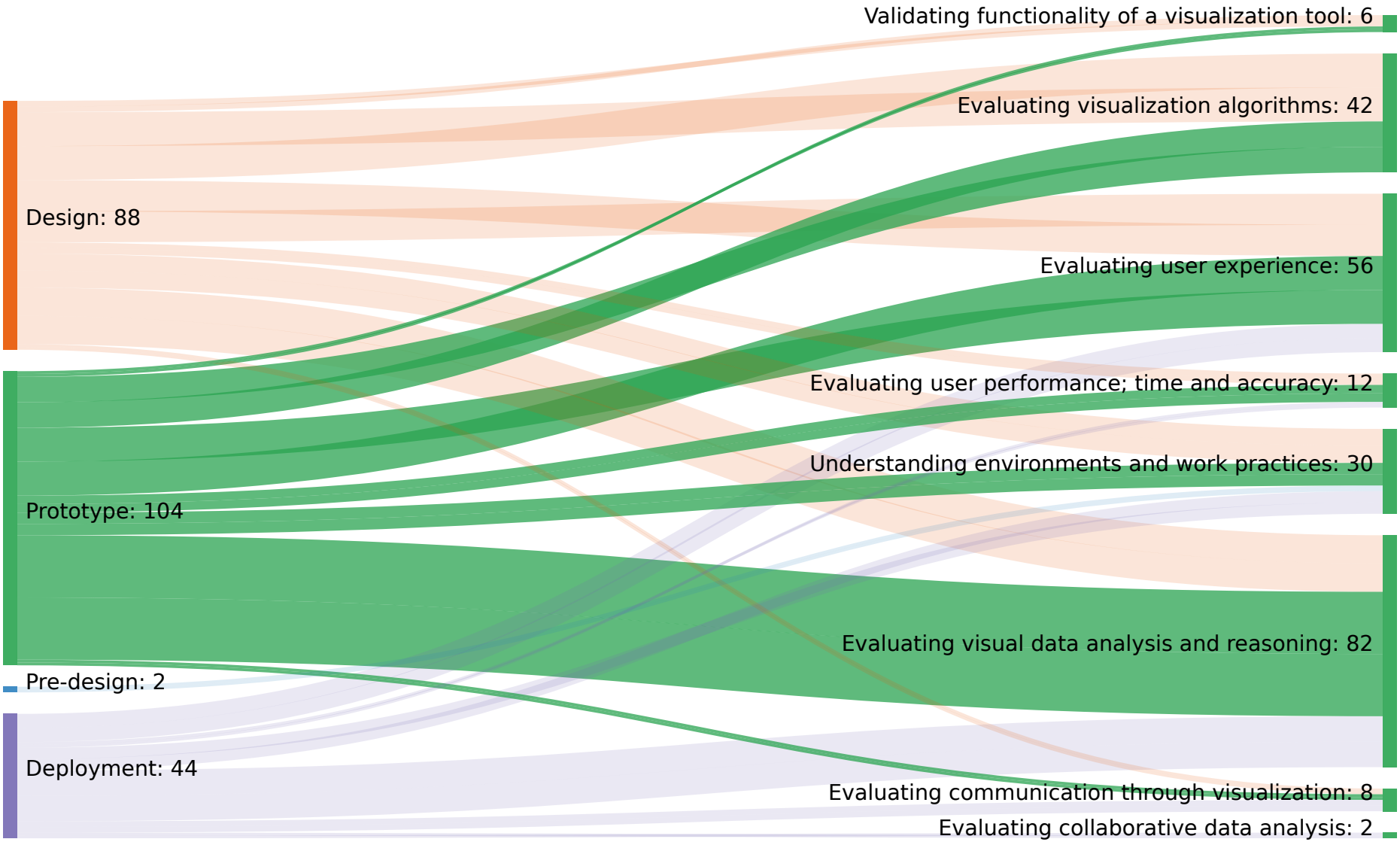
Evaluation scope vs. evaluation aspects



Evaluation aspects by Lam et al. (2012): *"Empirical Studies in Information Visualization: Seven Scenarios"*

What is evaluated, and when?

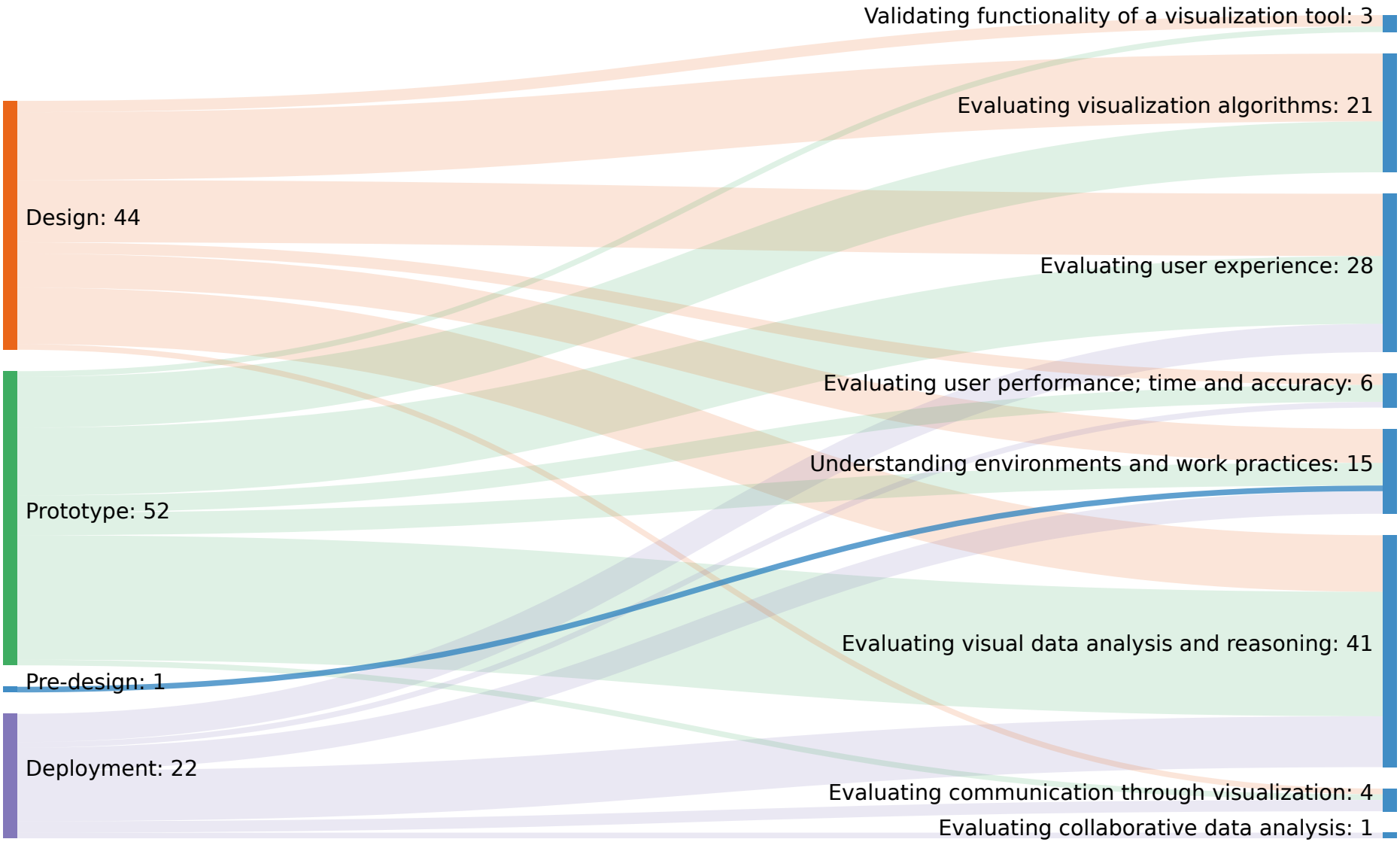
Evaluation scope vs. evaluation aspects



Evaluation aspects by Lam et al. (2012): *"Empirical Studies in Information Visualization: Seven Scenarios"*

What is evaluated, and when?

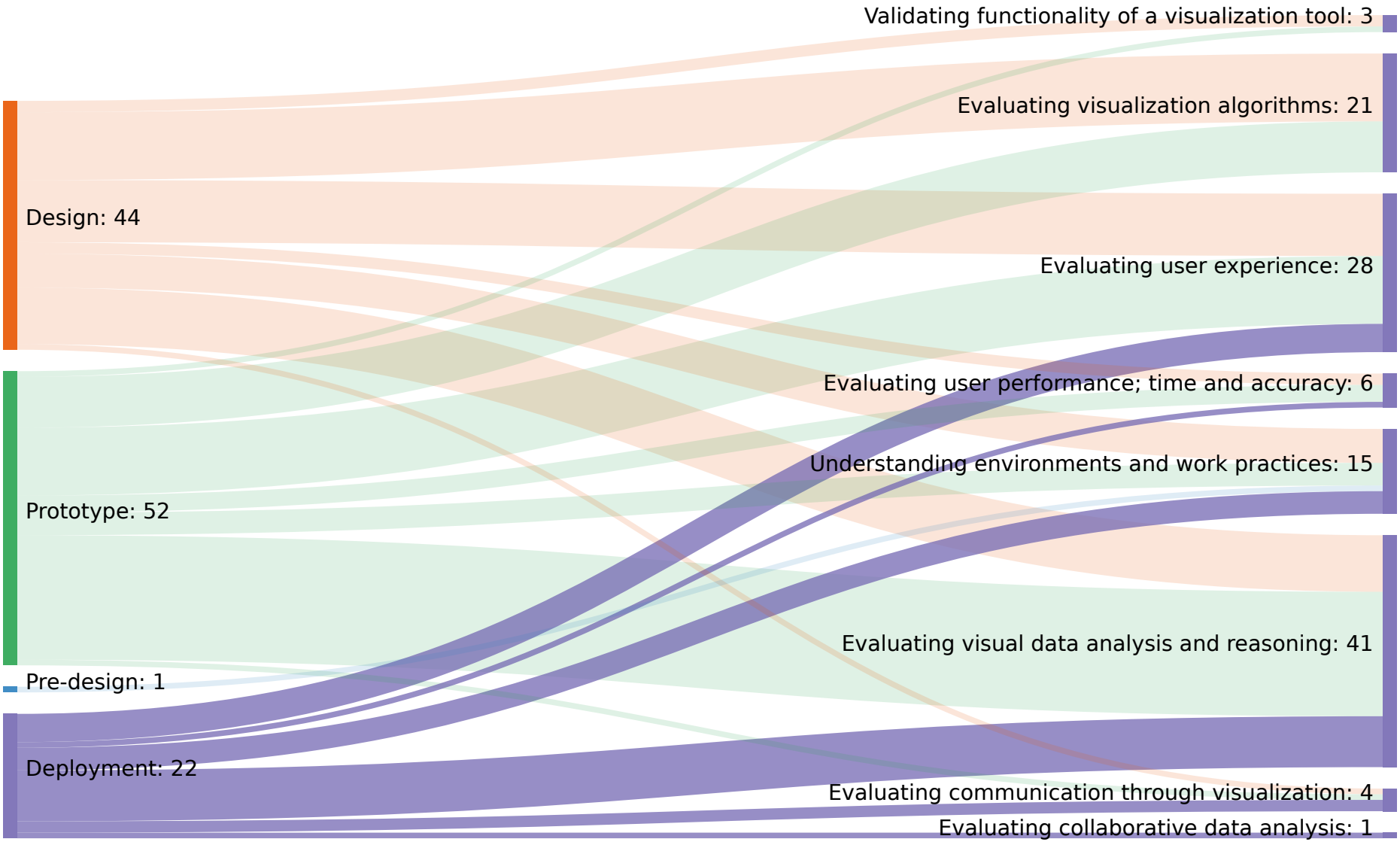
Evaluation scope vs. evaluation aspects



Evaluation aspects by Lam et al. (2012): *"Empirical Studies in Information Visualization: Seven Scenarios"*

What is evaluated, and when?

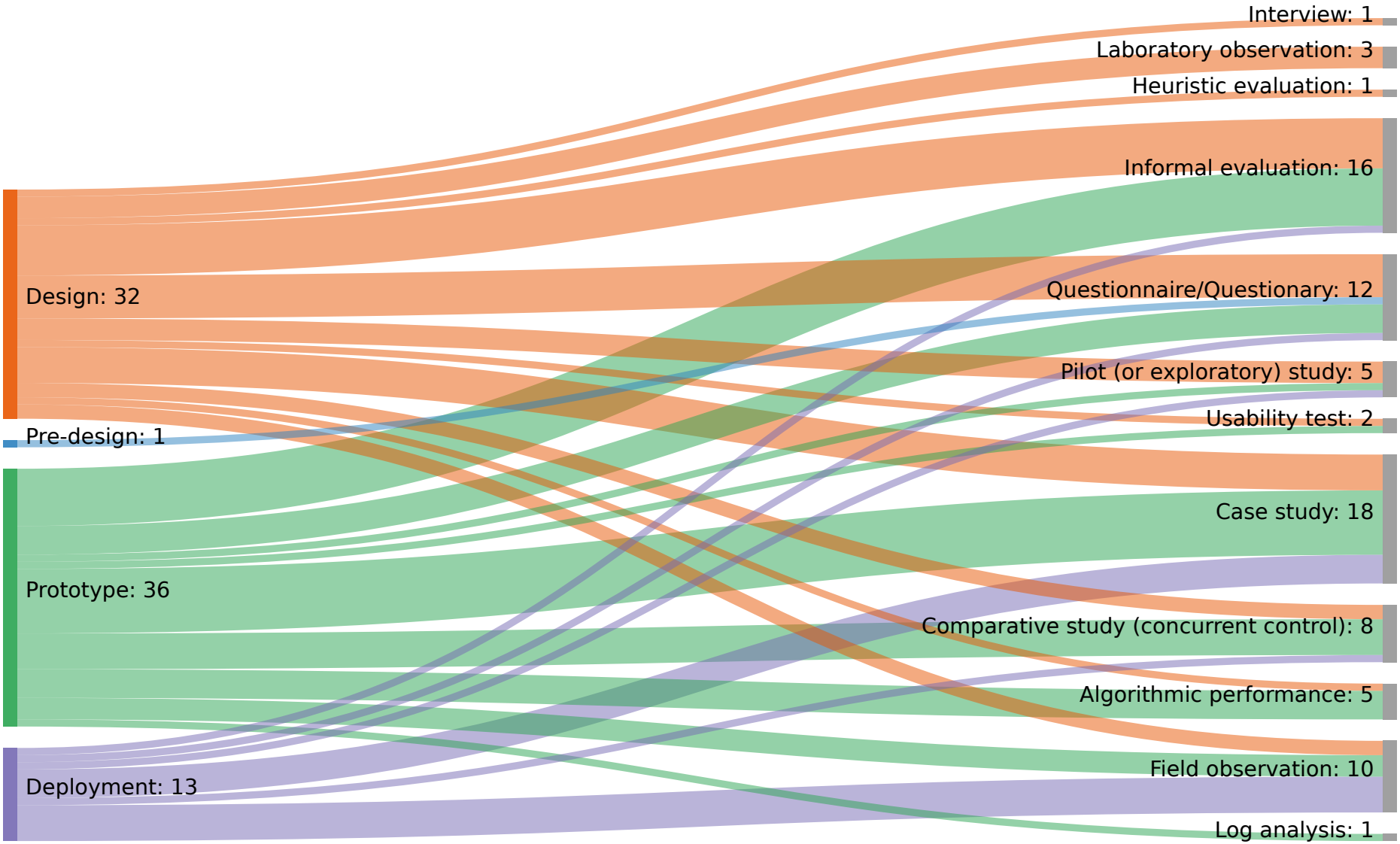
Evaluation scope vs. evaluation aspects



Evaluation aspects by Lam et al. (2012): *"Empirical Studies in Information Visualization: Seven Scenarios"*

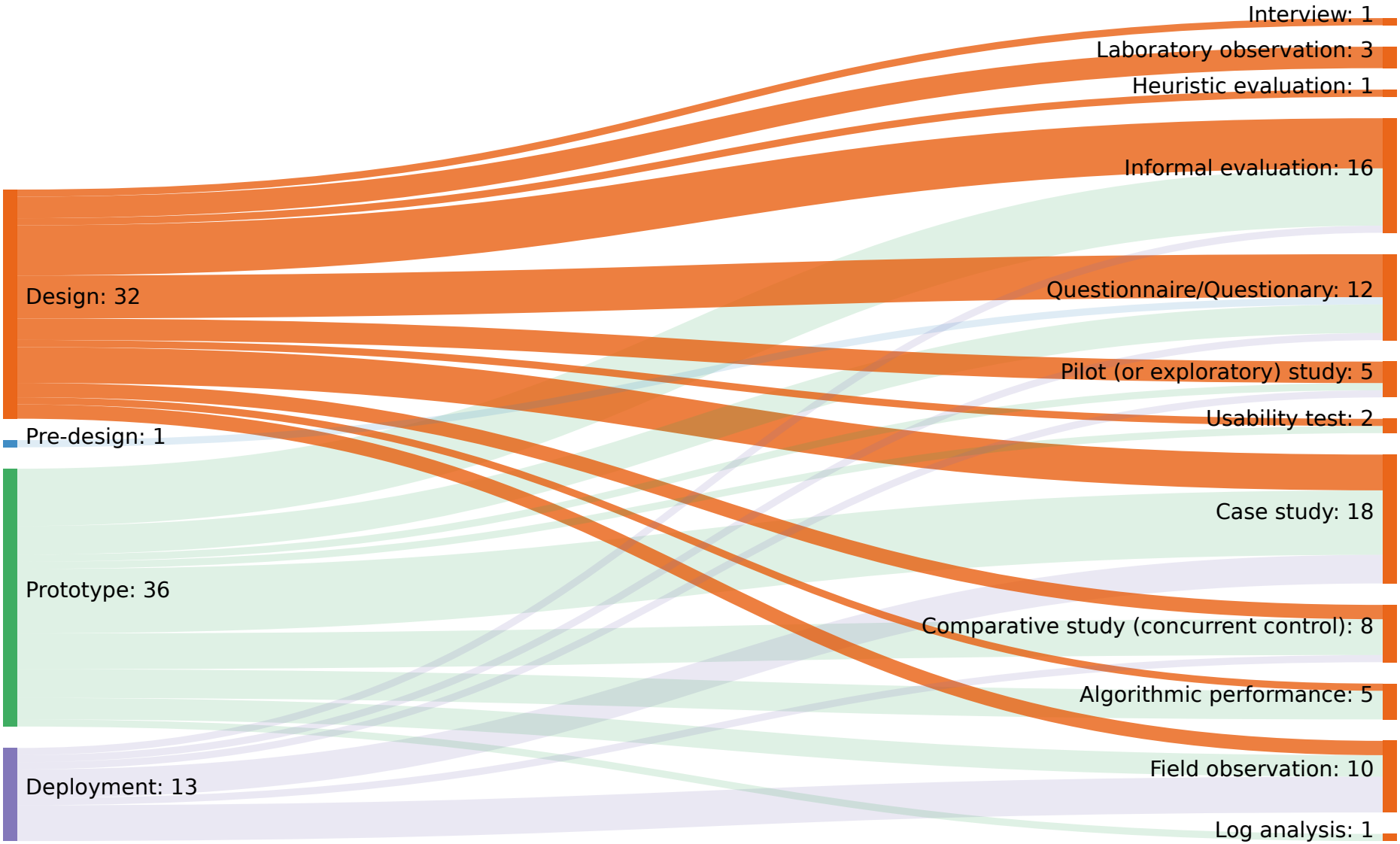
What is evaluated, and when?

Evaluation scope vs. evaluation methods



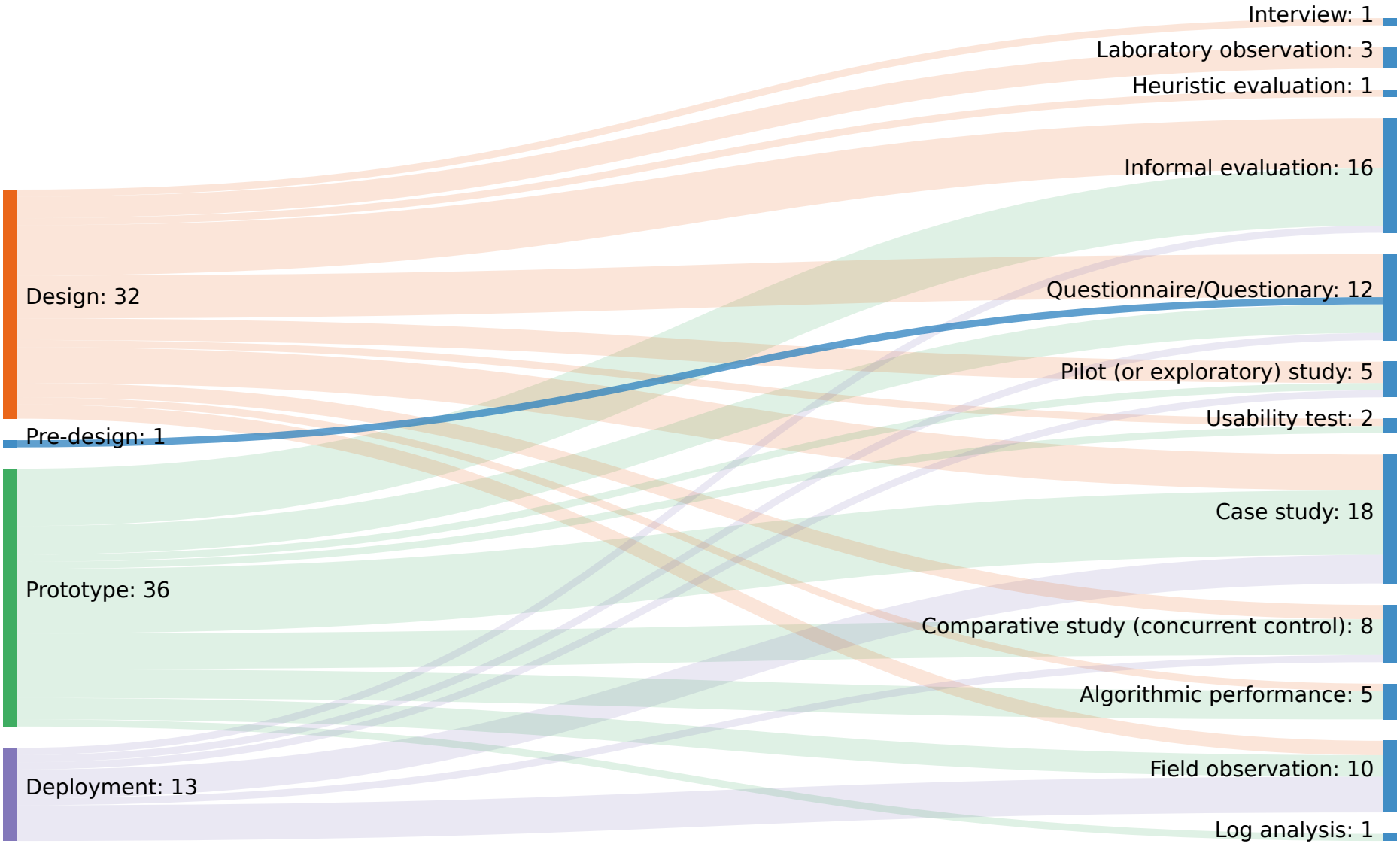
What is evaluated, and when?

Evaluation scope vs. evaluation methods



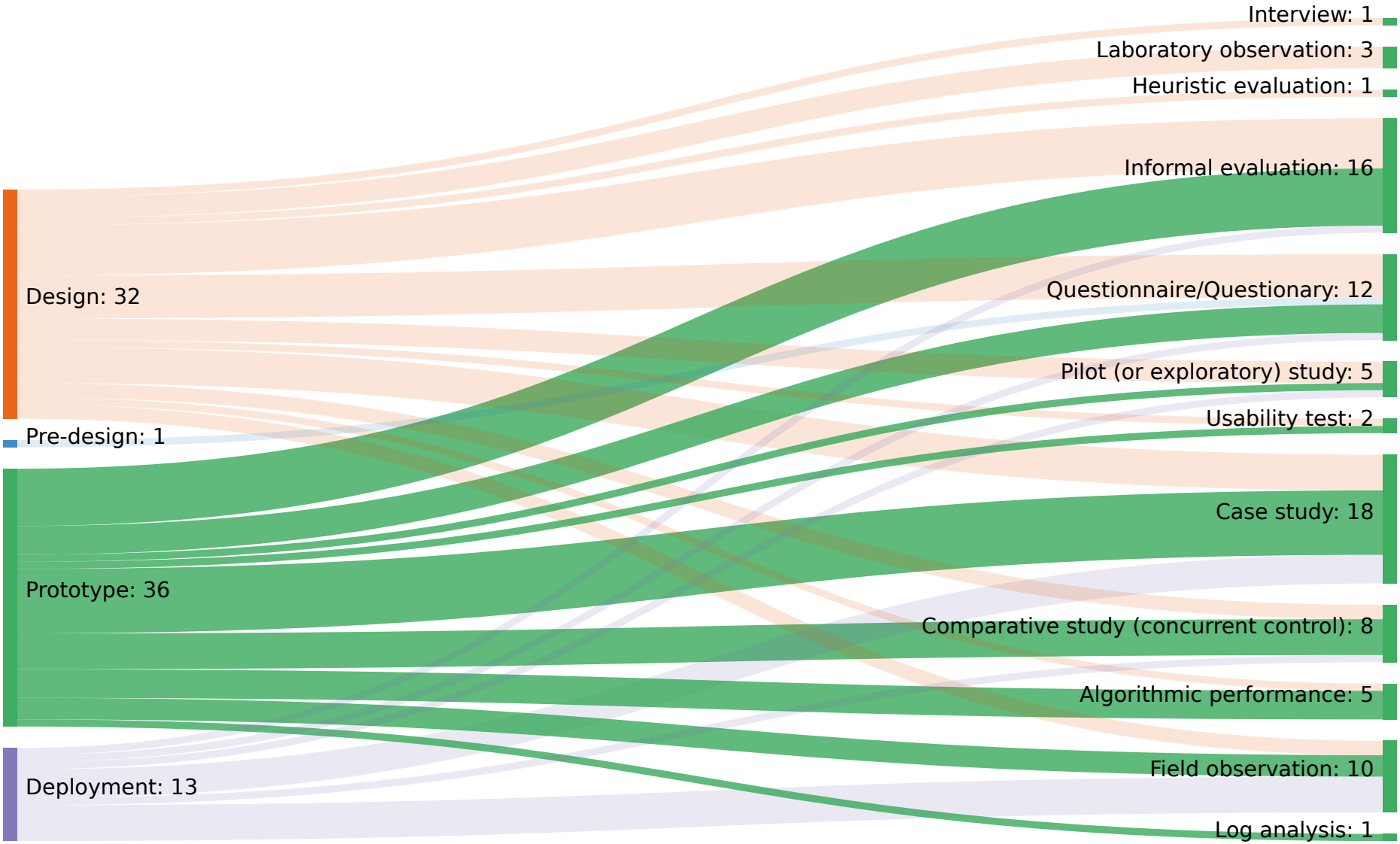
What is evaluated, and when?

Evaluation scope vs. evaluation methods



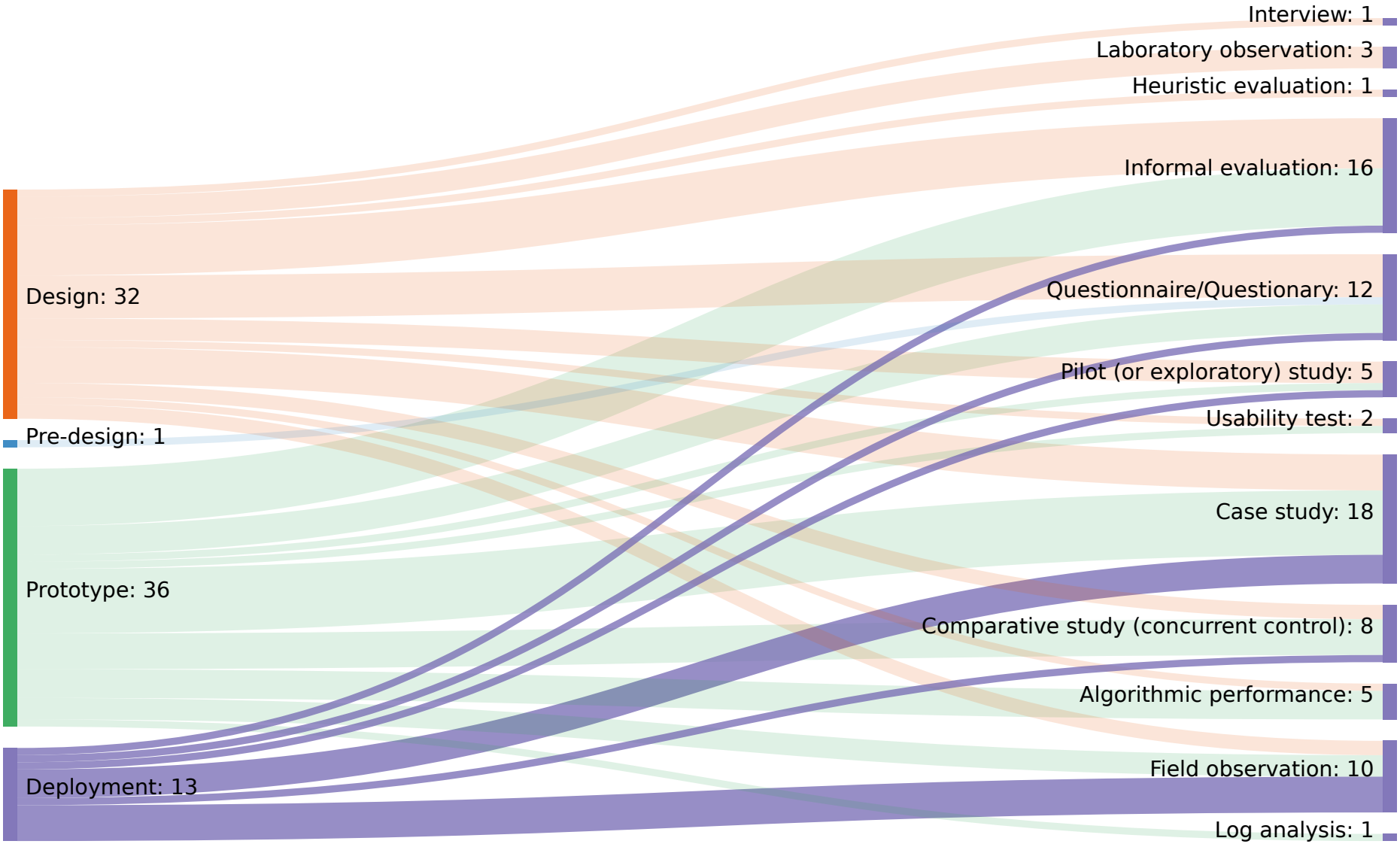
What is evaluated, and when?

Evaluation scope vs. evaluation methods



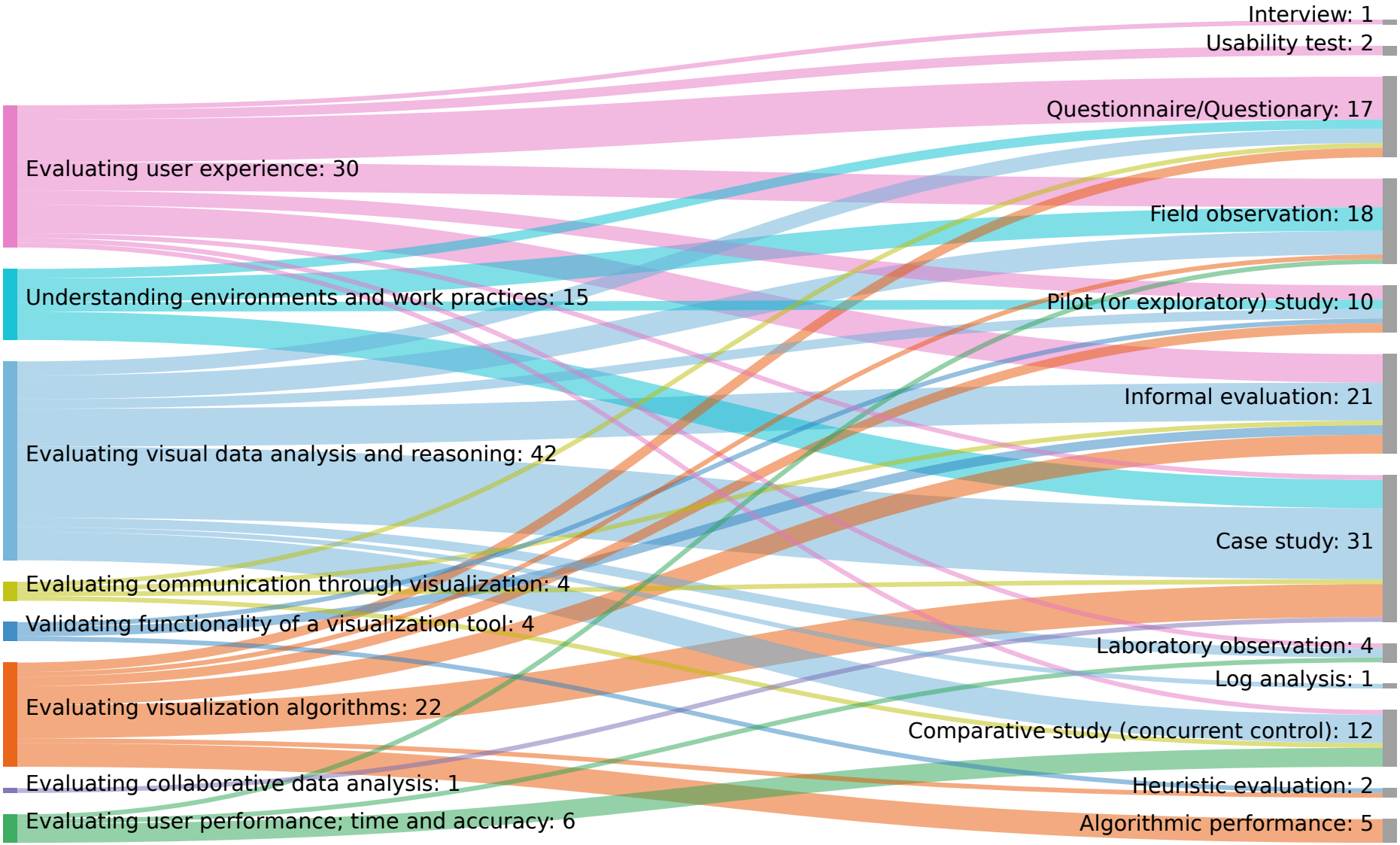
What is evaluated, and when?

Evaluation scope vs. evaluation methods



What is evaluated, and how?

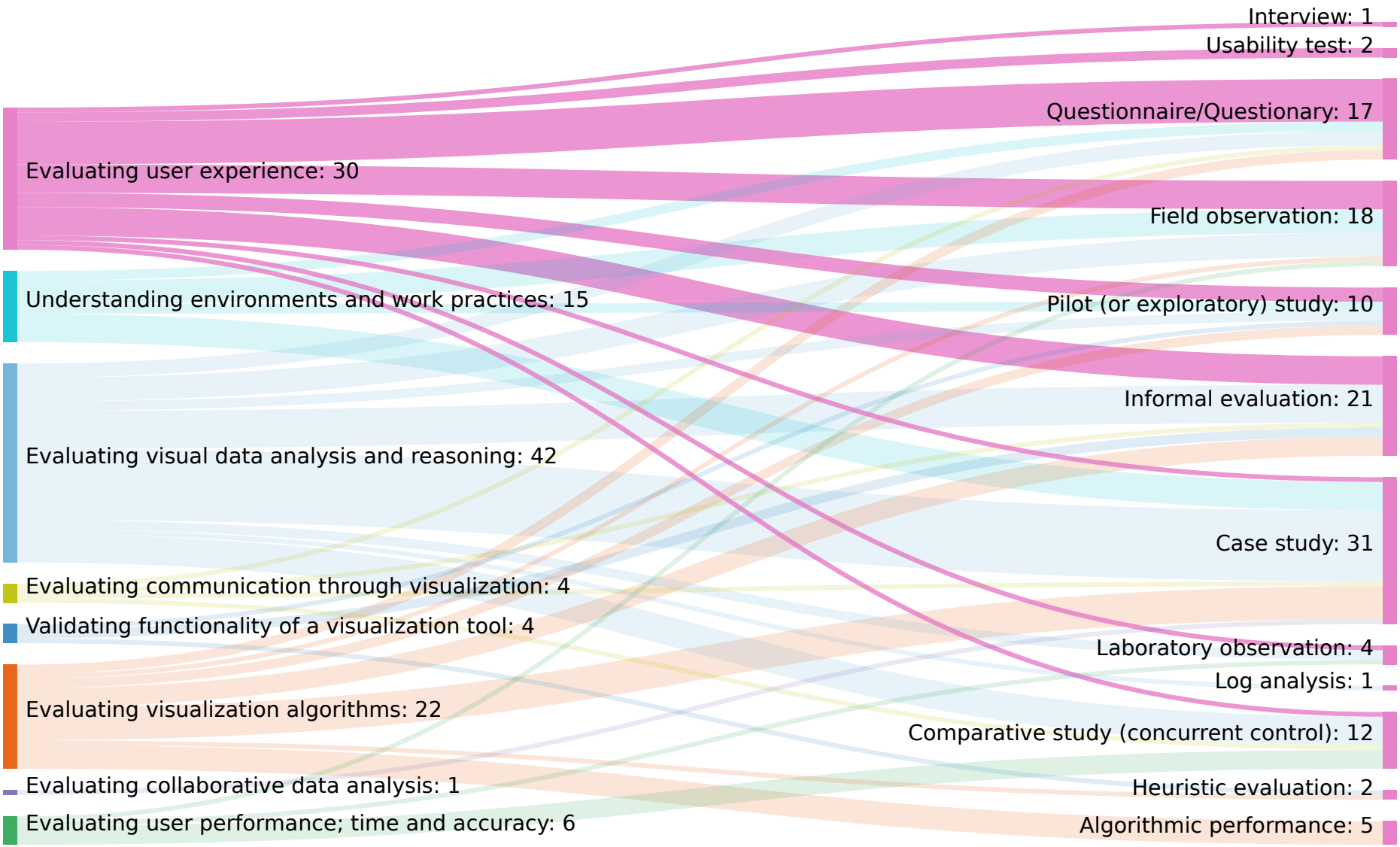
Evaluation aspects vs. evaluation methods



Evaluation aspects by Lam et al. (2012): *"Empirical Studies in Information Visualization: Seven Scenarios"*

What is evaluated, and how?

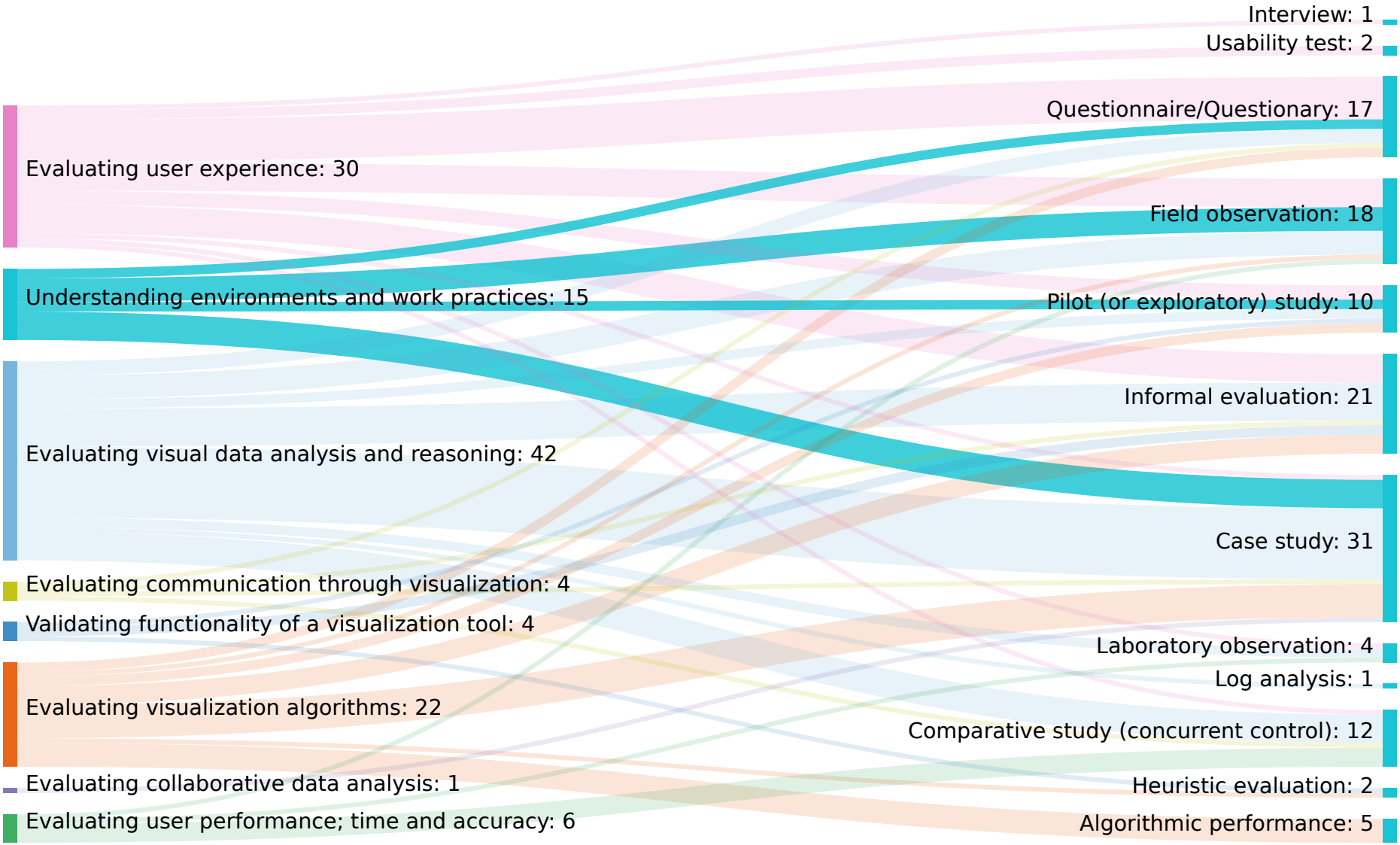
Evaluation aspects vs. evaluation methods



Evaluation aspects by Lam et al. (2012): *"Empirical Studies in Information Visualization: Seven Scenarios"*

What is evaluated, and how?

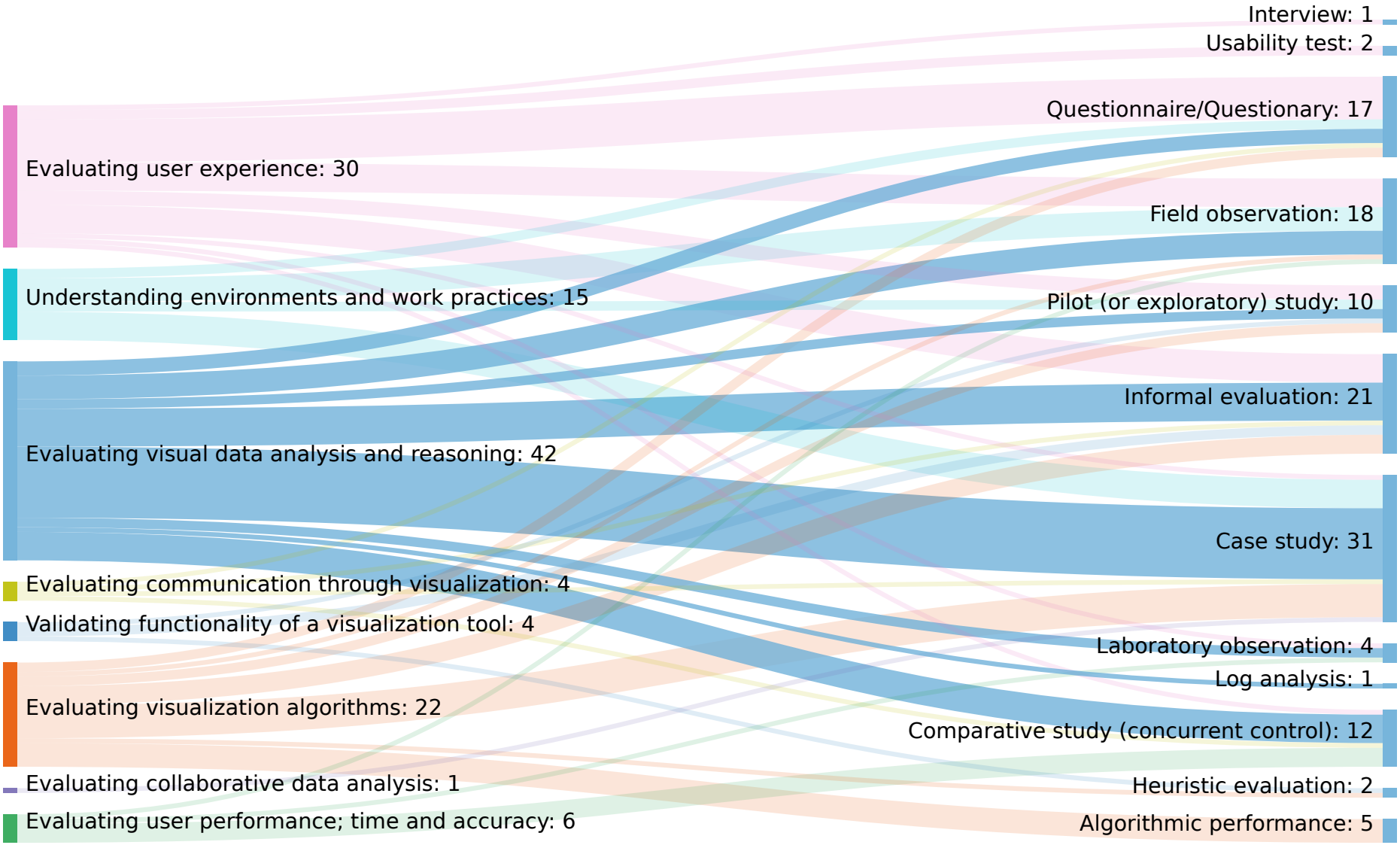
Evaluation aspects vs. evaluation methods



Evaluation aspects by Lam et al. (2012): *"Empirical Studies in Information Visualization: Seven Scenarios"*

What is evaluated, and how?

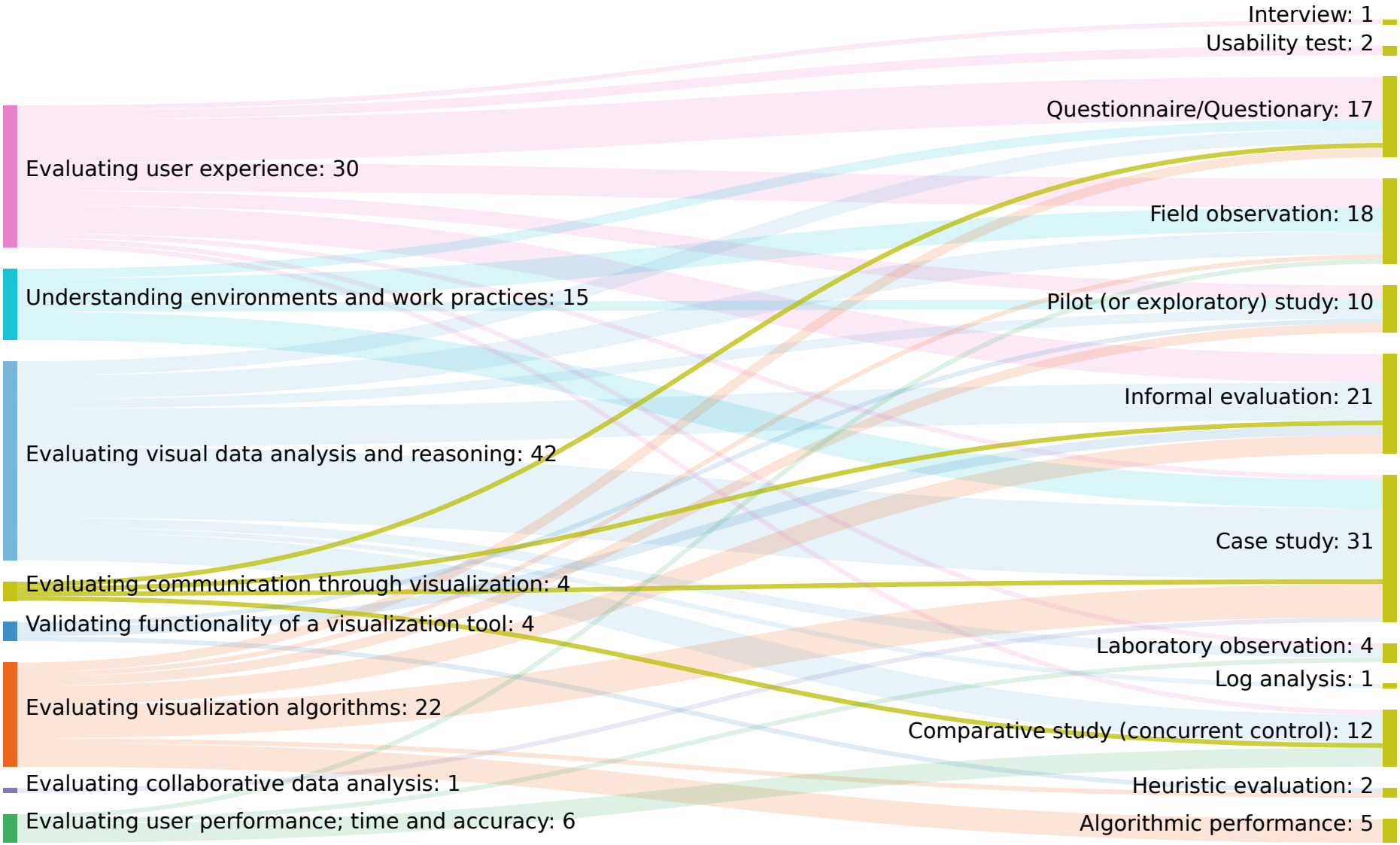
Evaluation aspects vs. evaluation methods



Evaluation aspects by Lam et al. (2012): *"Empirical Studies in Information Visualization: Seven Scenarios"*

What is evaluated, and how?

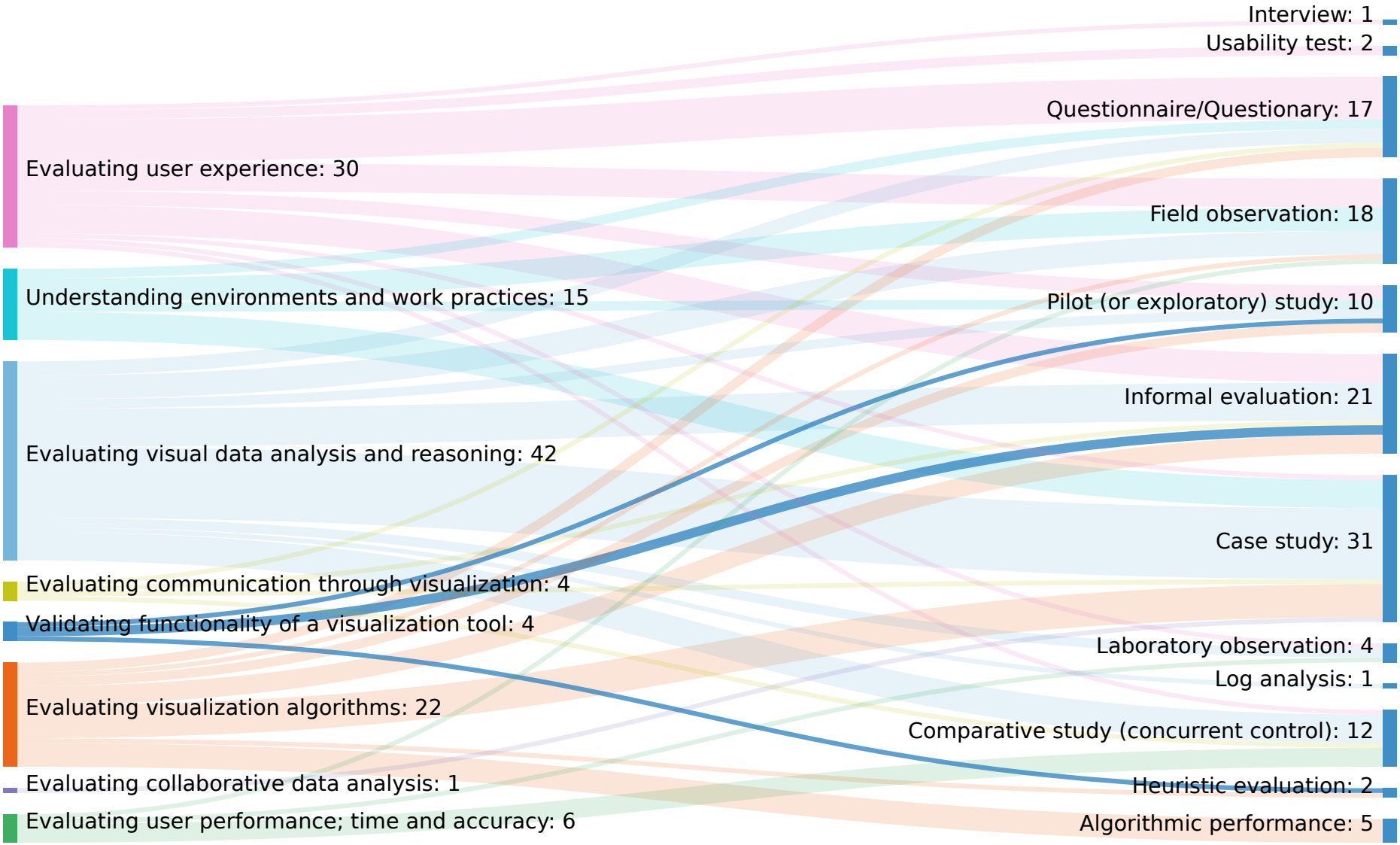
Evaluation aspects vs. evaluation methods



Evaluation aspects by Lam et al. (2012): *"Empirical Studies in Information Visualization: Seven Scenarios"*

What is evaluated, and how?

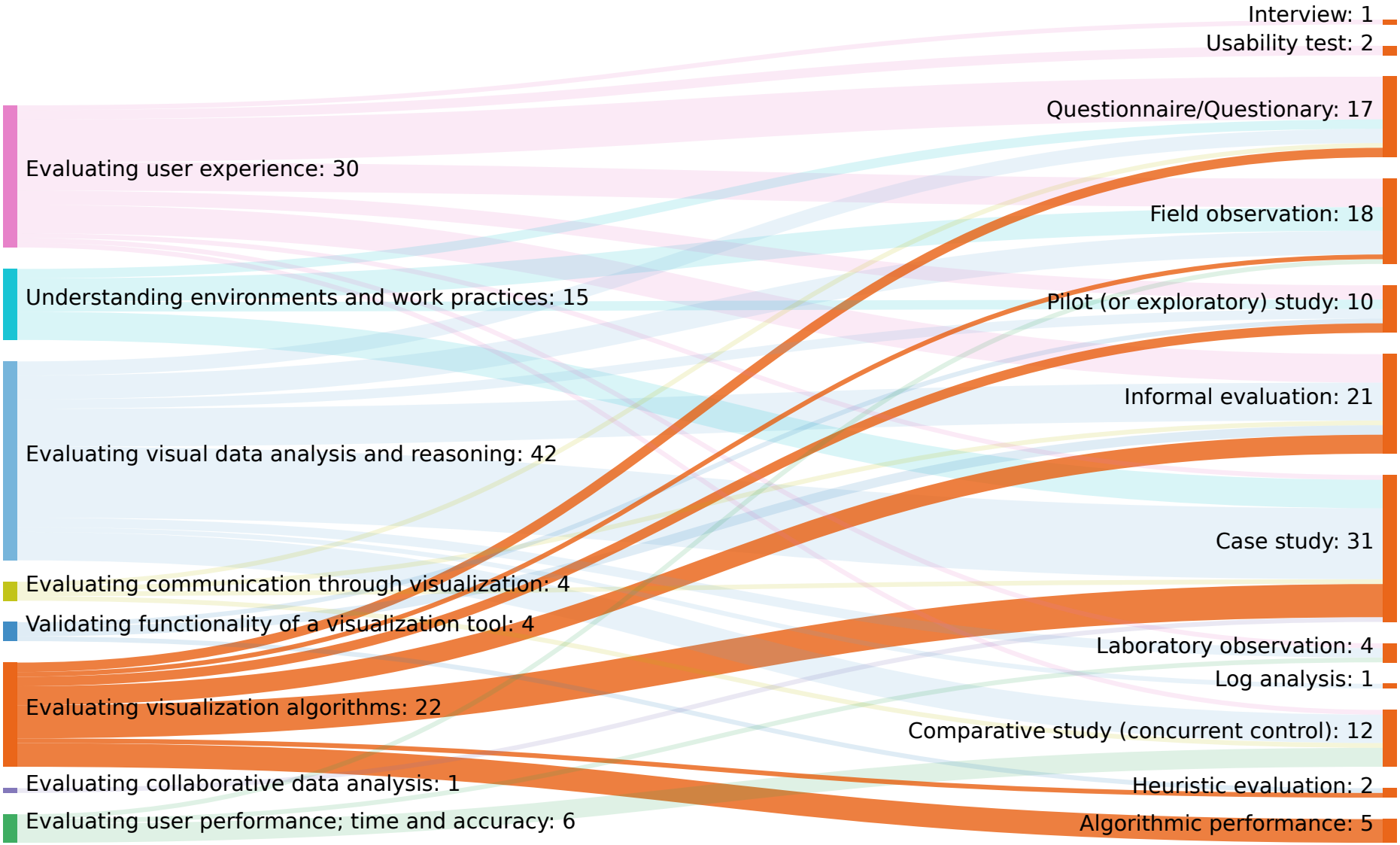
Evaluation aspects vs. evaluation methods



Evaluation aspects by Lam et al. (2012): *"Empirical Studies in Information Visualization: Seven Scenarios"*

What is evaluated, and how?

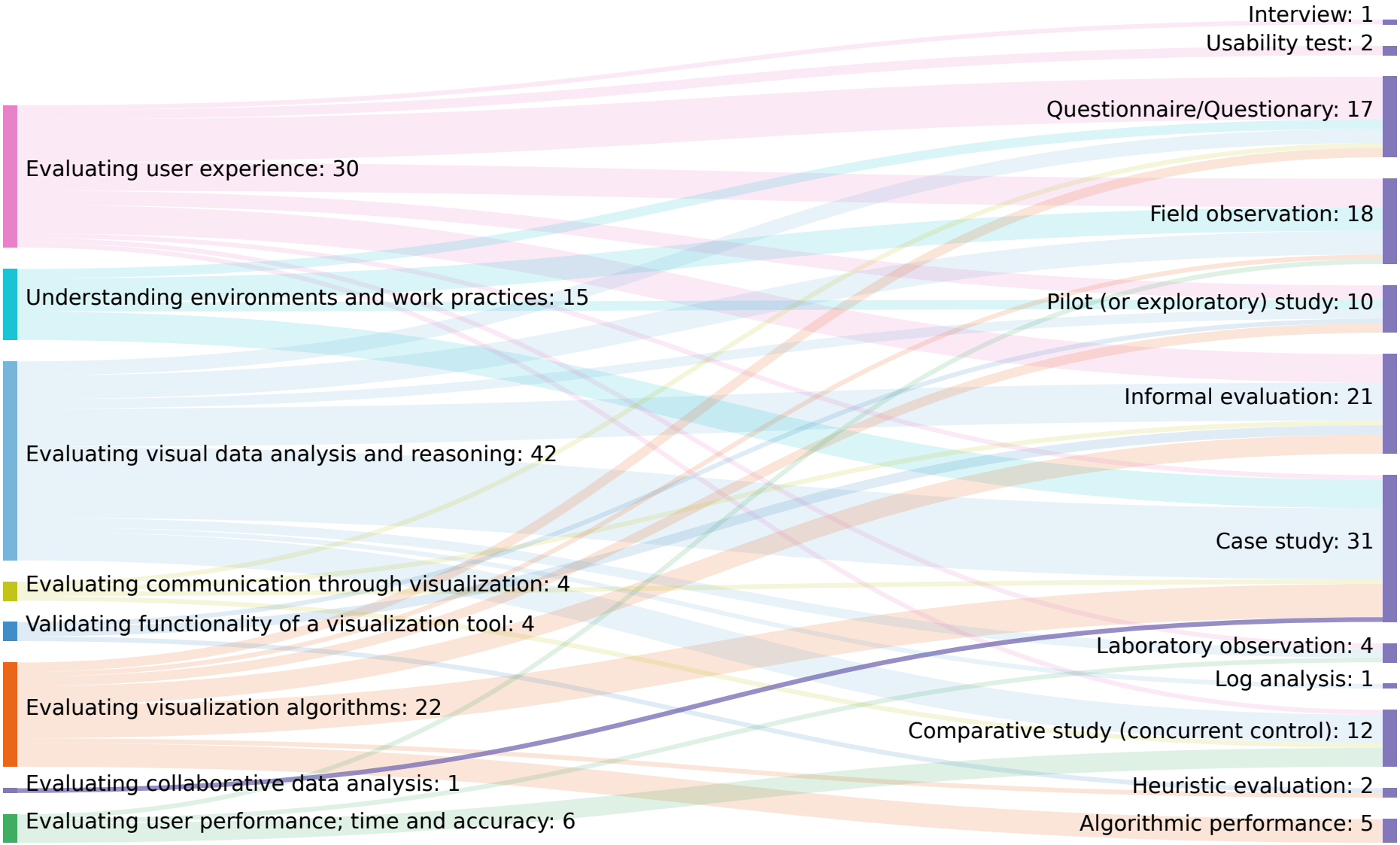
Evaluation aspects vs. evaluation methods



Evaluation aspects by Lam et al. (2012): *"Empirical Studies in Information Visualization: Seven Scenarios"*

What is evaluated, and how?

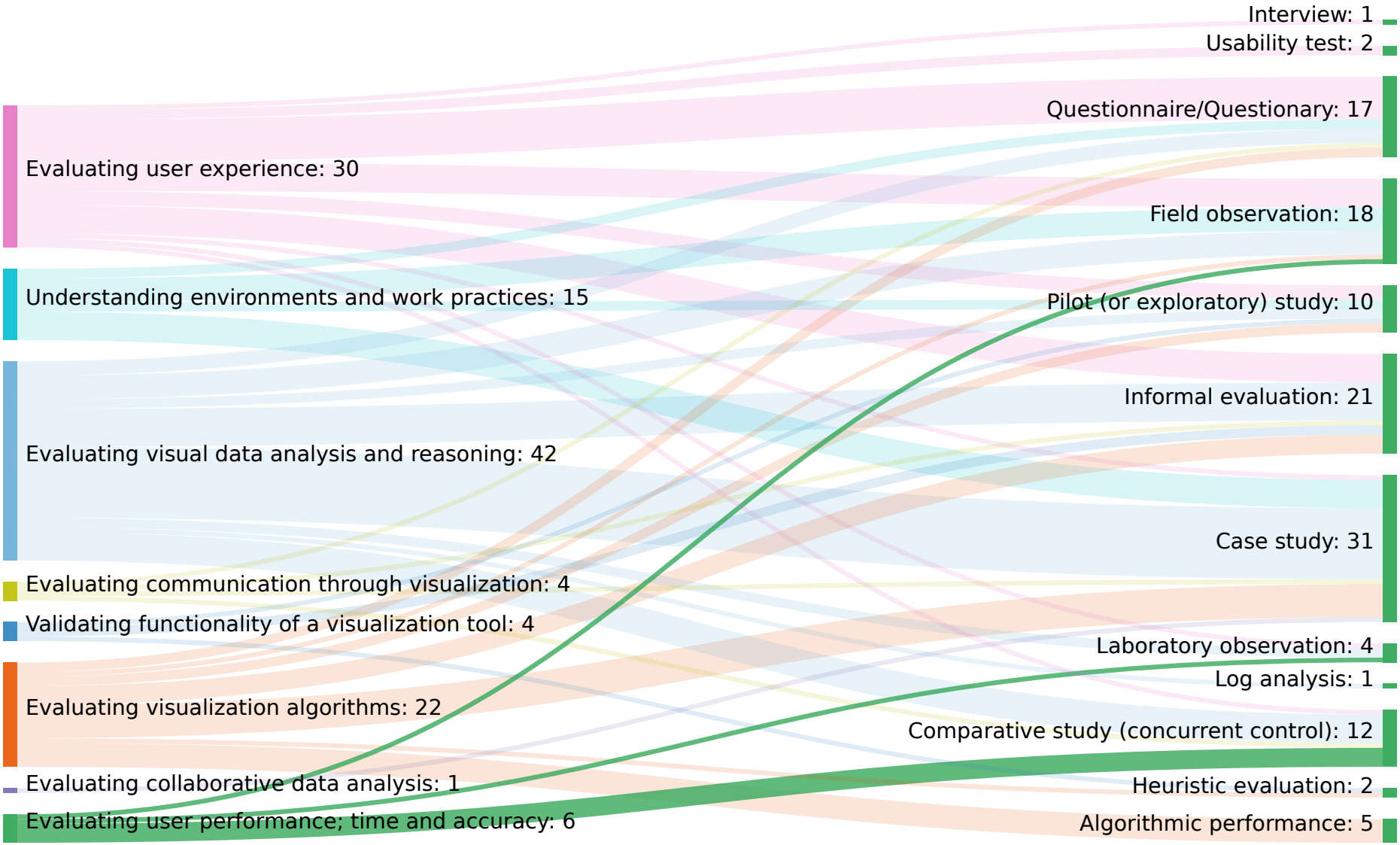
Evaluation aspects vs. evaluation methods



Evaluation aspects by Lam et al. (2012): *"Empirical Studies in Information Visualization: Seven Scenarios"*

What is evaluated, and how?

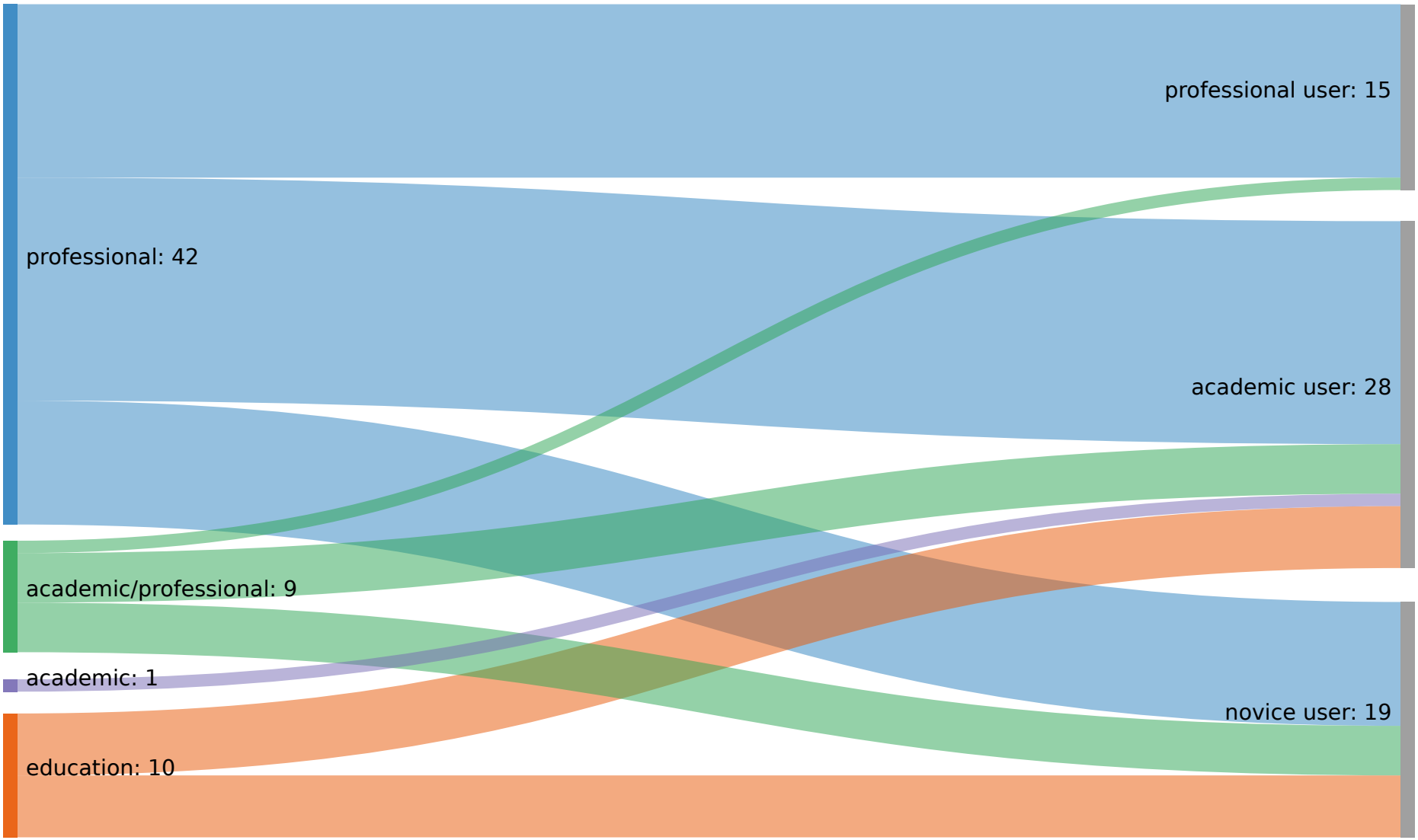
Evaluation aspects vs. evaluation methods



Evaluation aspects by Lam et al. (2012): *"Empirical Studies in Information Visualization: Seven Scenarios"*

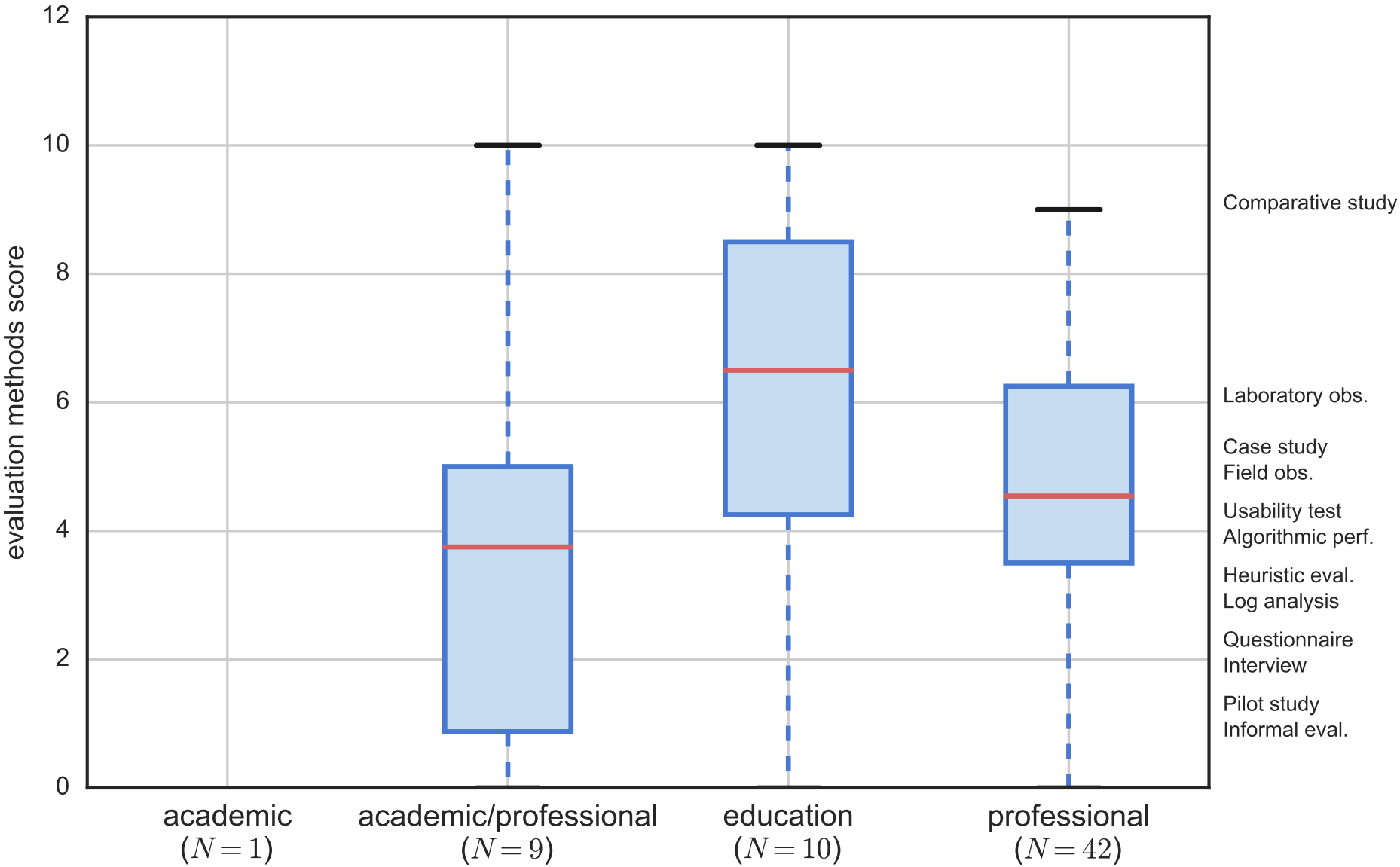
Who are the evaluation subjects?

Target audience vs. evaluation subjects



And the winner is. . .

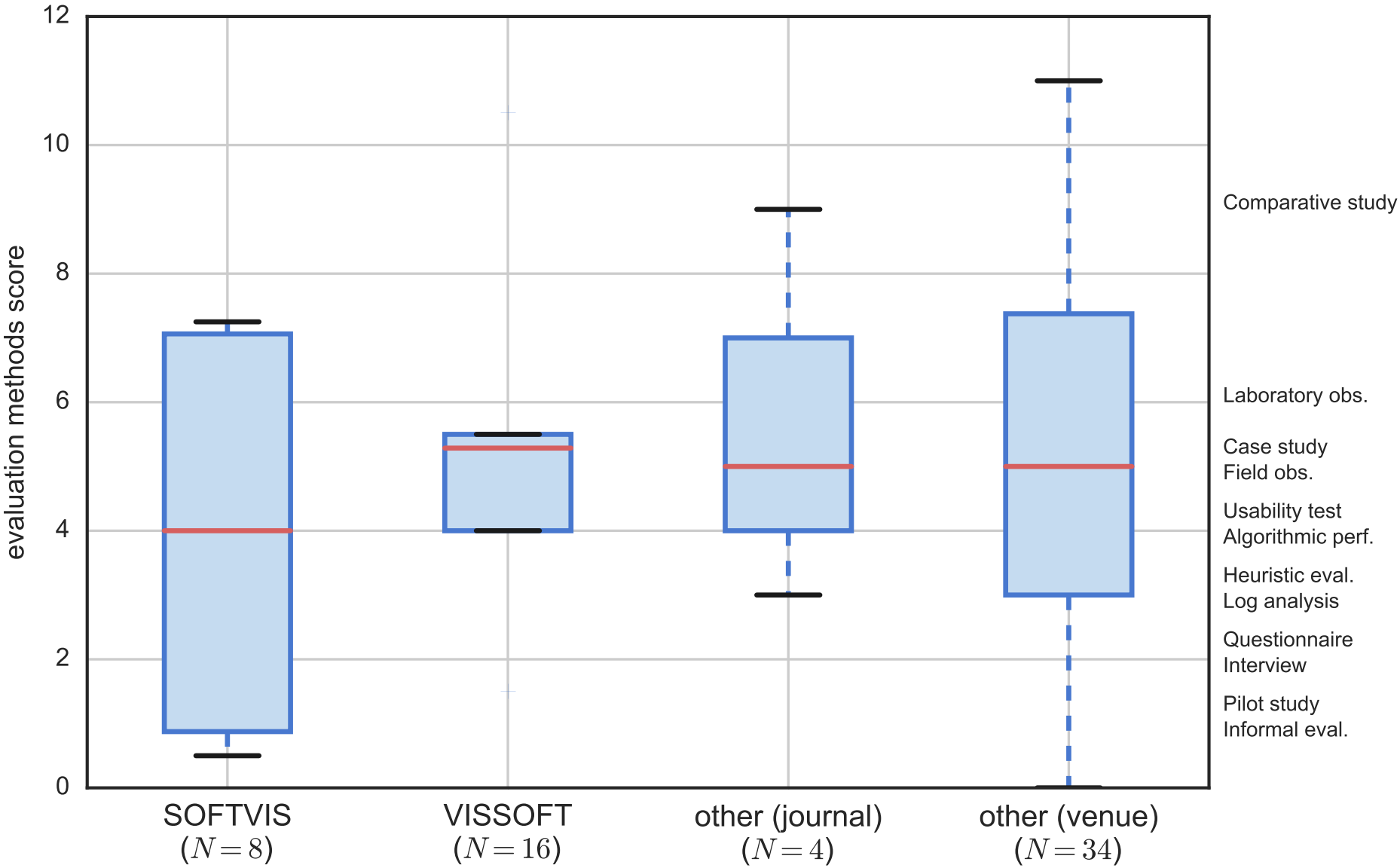
Evaluation score by target audience



Evaluation method ranking roughly based on B. Kitchenham's *"Study design hierarchy for Software Engineering"*

And the winner is. . .

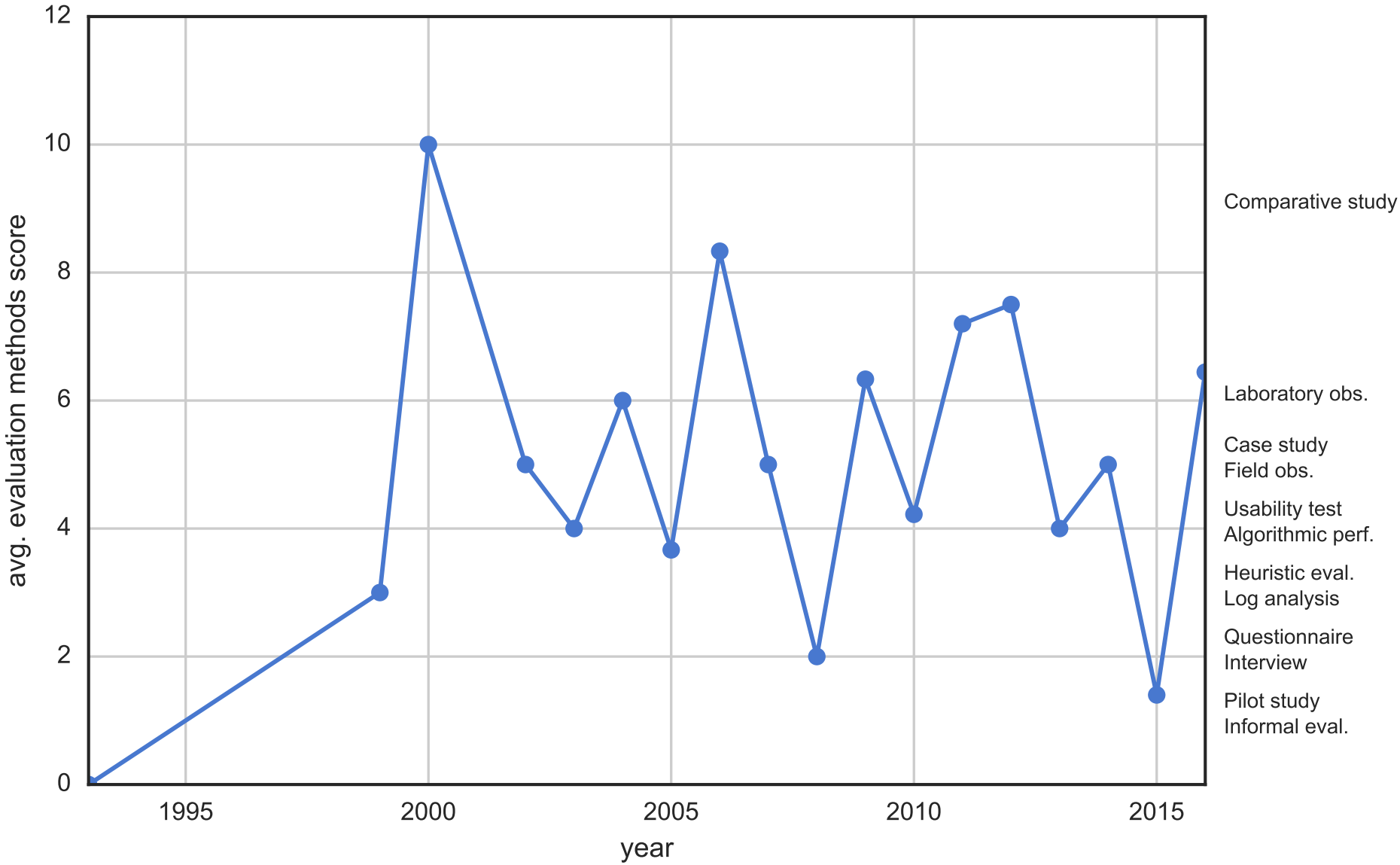
Evaluation score by venue



Evaluation method ranking roughly based on B. Kitchenham's *"Study design hierarchy for Software Engineering"*

Does the evaluation score improve over time?

Evaluation score by year



Evaluation method ranking roughly based on B. Kitchenham's *"Study design hierarchy for Software Engineering"*

Discussion

Results

Discussion

- Widely different software visualization needs based on:
  - the target audience (e.g. professional vs. education)
  - the programming language/paradigm

Results

Discussion

- Widely different software visualization needs based on:
  - the target audience (e.g. professional vs. education)
  - the programming language/paradigm
- Naturally different evaluation aspects and methods depending on evaluation scope (or stage):
  - Informal evaluation methods in early evaluation scopes
  - More formal evaluation methods with later evaluation scopes
  - Overall rather low evaluation effort
  - Minimum standard seems to have been established

Results

Discussion

- Widely different software visualization needs based on:
  - the target audience (e.g. professional vs. education)
  - the programming language/paradigm
- Naturally different evaluation aspects and methods depending on evaluation scope (or stage):
  - Informal evaluation methods in early evaluation scopes
  - More formal evaluation methods with later evaluation scopes
  - Overall rather low evaluation effort
  - Minimum standard seems to have been established
- Consolidation from other venues to SOFTVIS/VISSOFT

Limitations/threats to validity

Discussion

- Selection bias (ACM, IEEE, SD)

Limitations/threats to validity

Discussion

- Selection bias (ACM, IEEE, SD)
- Scope of a seminar project vs. systematic literature review
  - Iterative process (expertise and time)
  - Establishing consensus (peer-review/repeatable?)

Limitations/threats to validity

Discussion

- Selection bias (ACM, IEEE, SD)
- Scope of a seminar project vs. systematic literature review
  - Iterative process (expertise and time)
  - Establishing consensus (peer-review/repeatable?)
- Paper classification (What is a design study?)

Limitations/threats to validity

Discussion

- Selection bias (ACM, IEEE, SD)
- Scope of a seminar project vs. systematic literature review
  - Iterative process (expertise and time)
  - Establishing consensus (peer-review/repeatable?)
- Paper classification (What is a design study?)
- Data extraction

Thanks for listening!

Thanks for listening! Questions?