Merging Directly-Follows Graphs and Sankey Diagrams for Visualizing Acyclic Processes

Ilya D. Derezovskiy, Nikita D. Shaimov, Irina A. Lomazova, Alexey A. Mitsyuk

HSE University, Faculty of Computer Science, PAIS Lab
Moscow, Russia

iderezovskiy@hse.ru, nshaimov@hse.ru, ilomazova@hse.ru, amitsyuk@hse.ru

SYRCoSE 2024
May 30, 2024

Introduction

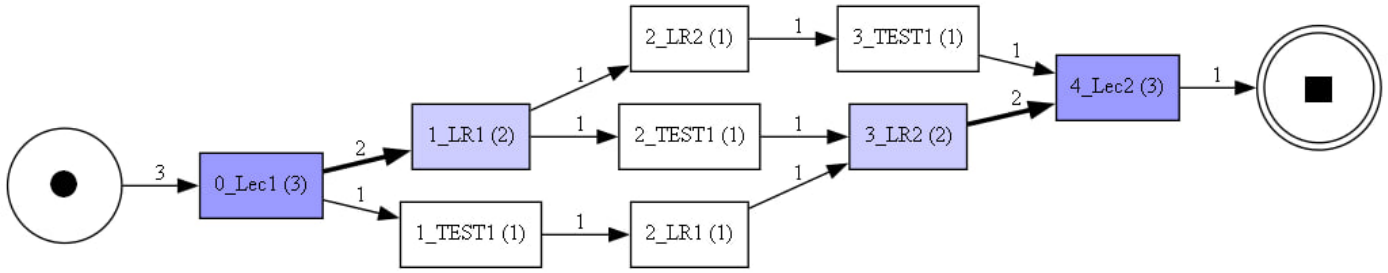


Fig. 1 An example of a DFG model

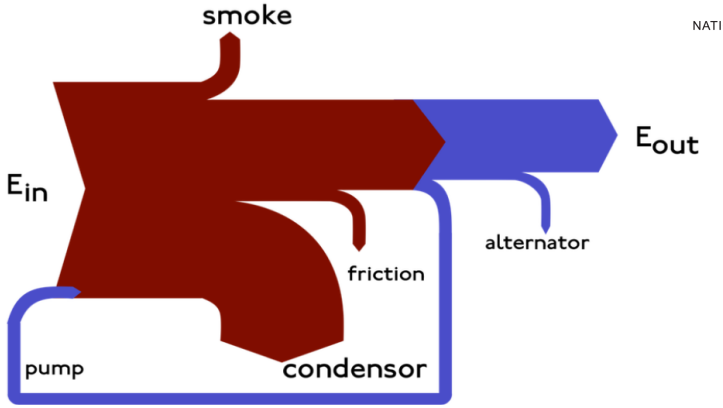


Fig. 2 An example of a Sankey diagram

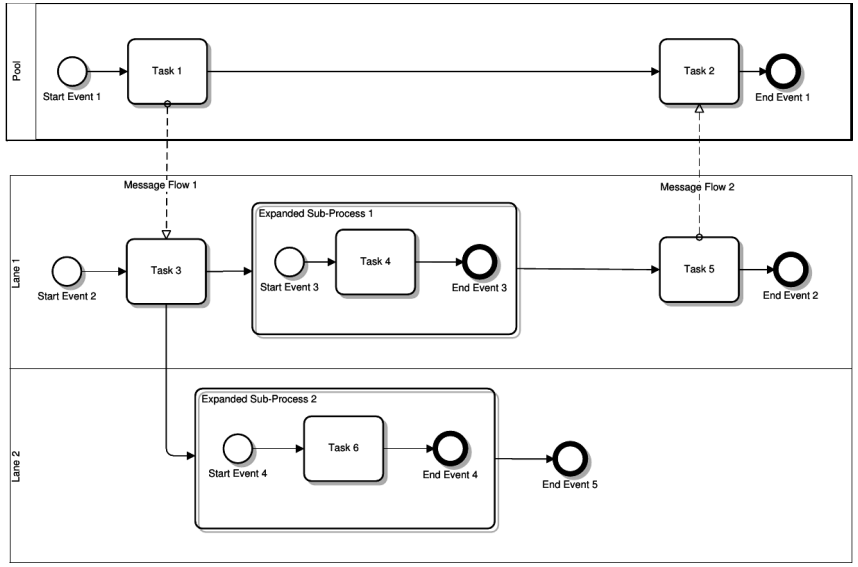


Fig. 3 An example of a BPMN model

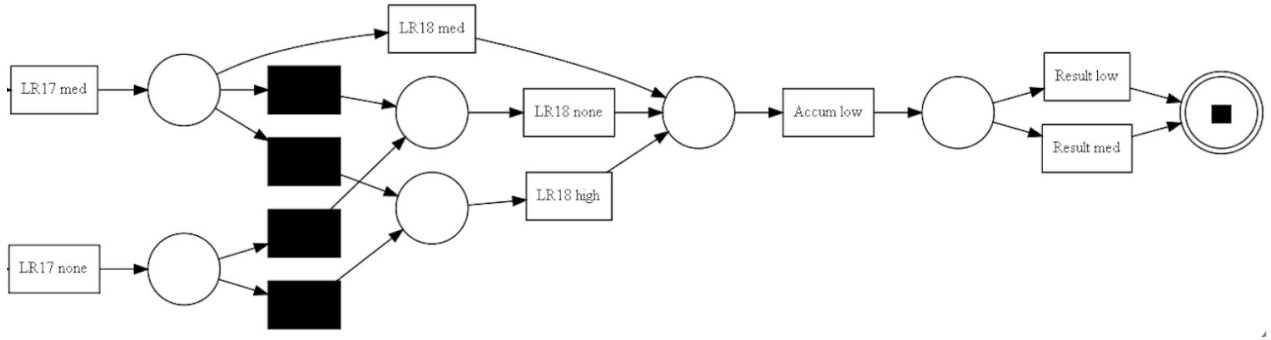


Fig. 4 An example of a Petri Net model

Overview of existing solutions

- PMTK (novel web-based Process Mining ToolKit)

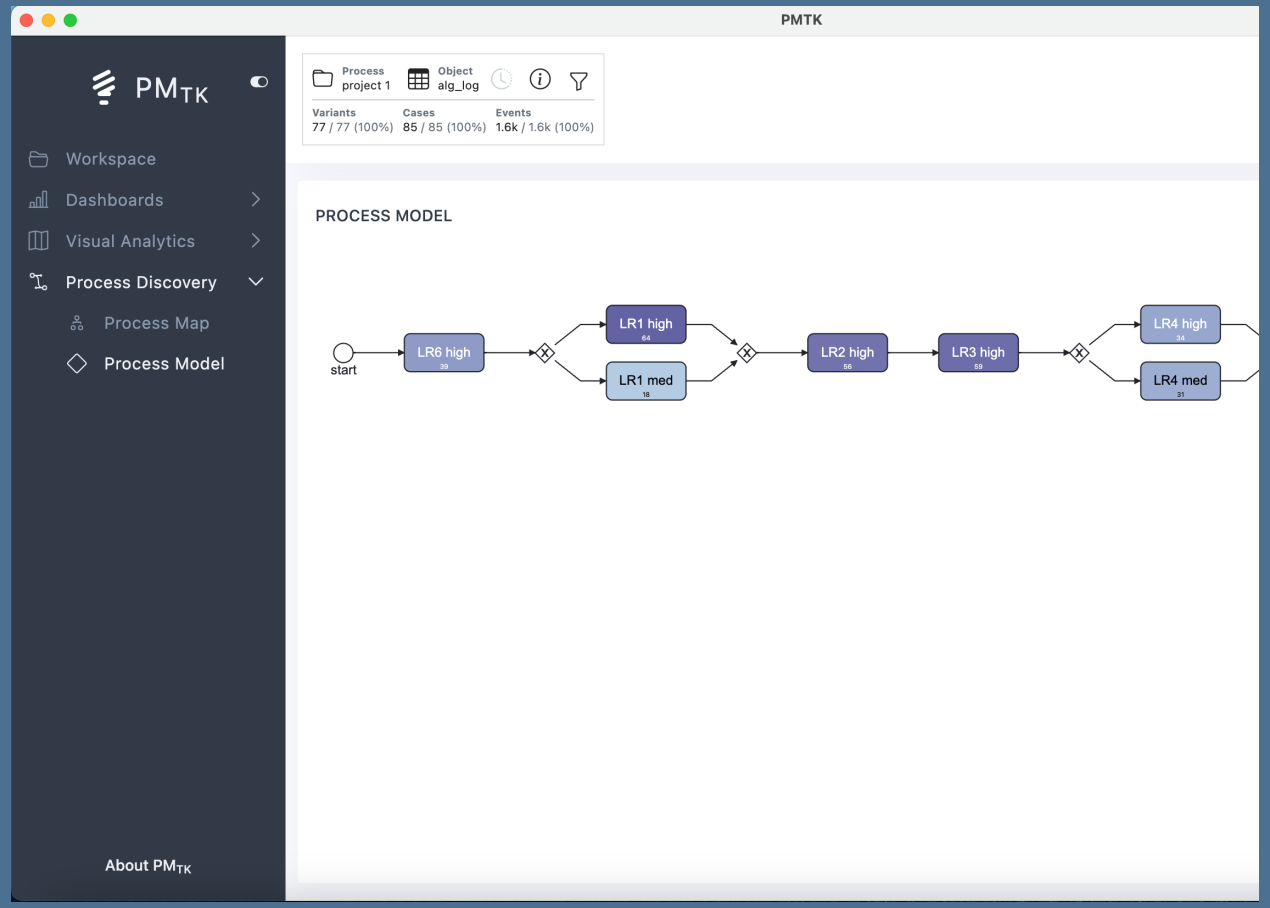


Fig. 5 The interface of PMTK application

Overview of existing solutions

- PMTK (novel web-based Process Mining ToolKit)
- **Celonis (commercial online platform for analysis business processes)**

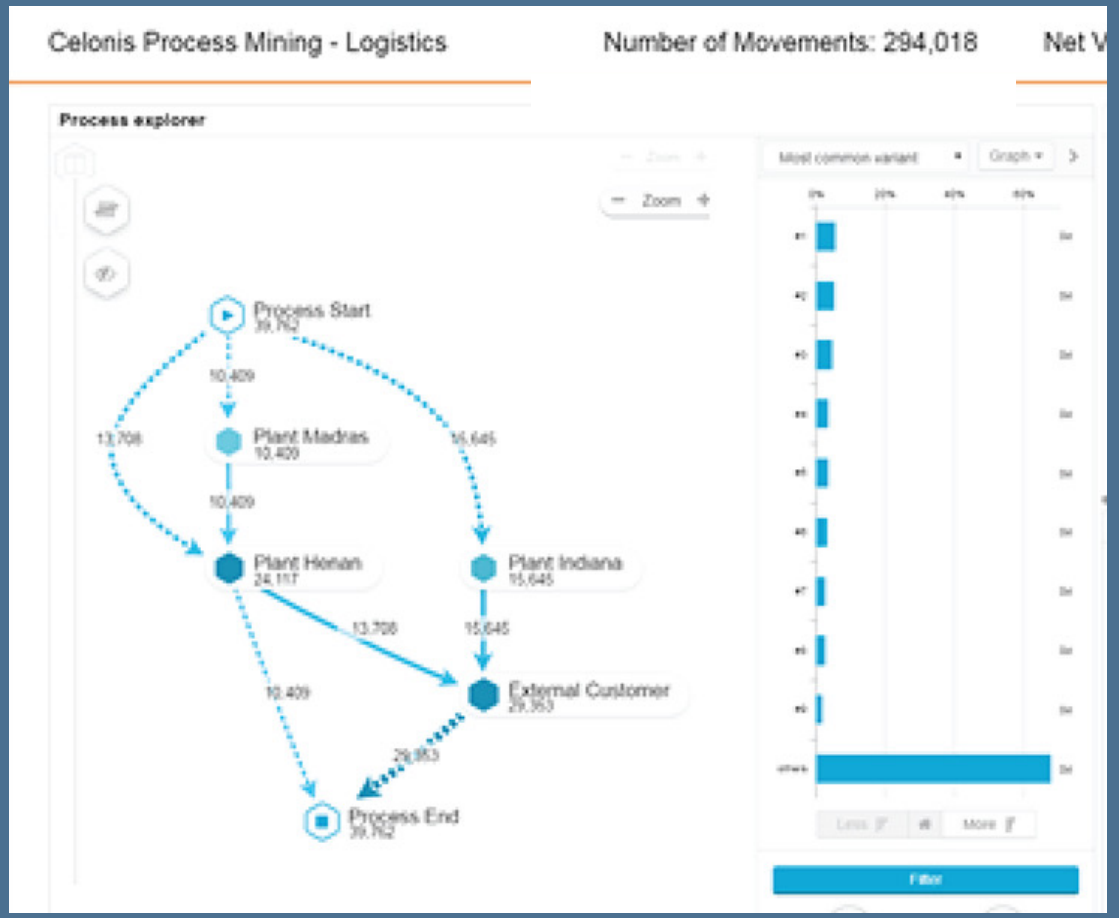


Fig. 6 The interface of Celonis application

Overview of existing solutions

- PMTK (novel web-based Process Mining ToolKit)
- Celonis (commercial online platform for analysis business processes)
- **SankeyMATIC (website for building Sankey diagrams)**

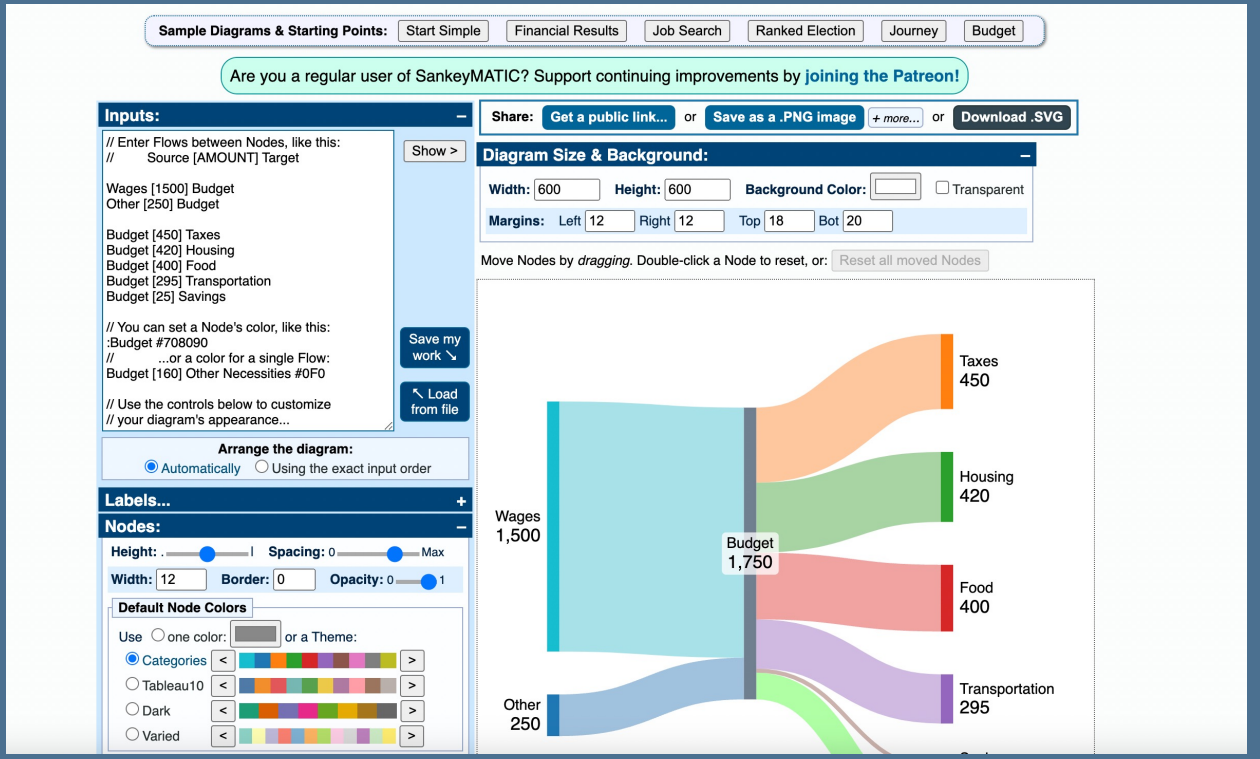


Fig. 7 The interface of SankeyMATIC application

Overview of existing solutions

- PMTK (novel web-based Process Mining ToolKit)
- Celonis (commercial online platform for analysis business processes)
- SankeyMATIC (website for building Sankey diagrams)
- **Fluxicon Disco (application for analysis business processes)**

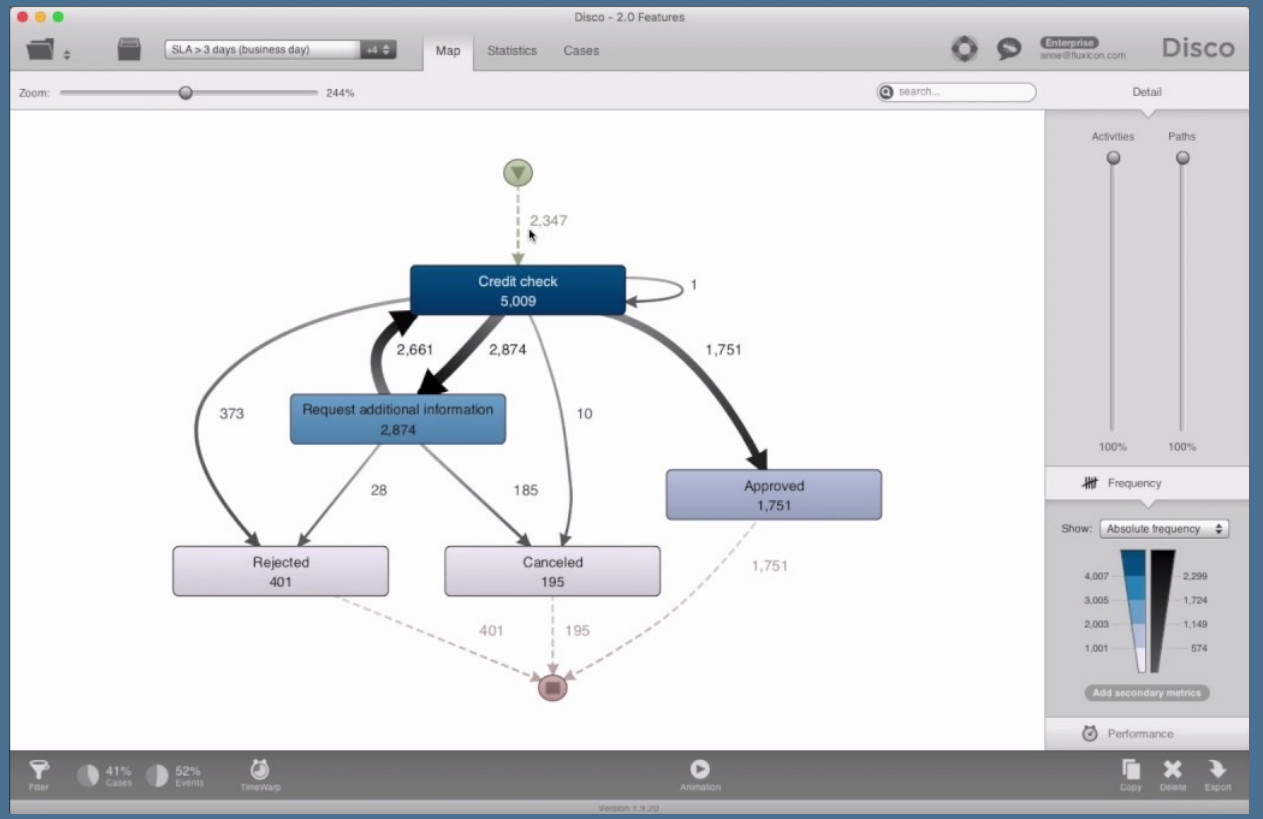


Fig. 8 The interface of Disco application

Overview of existing solutions

- PMTK (novel web-based Process Mining ToolKit)
- Celonis (commercial online platform for analysis business processes)
- SankeyMATIC (website for building Sankey diagrams)
- Fluxicon Disco (application for analysis business processes)
- **ProM (plugin-based framework for process mining)**

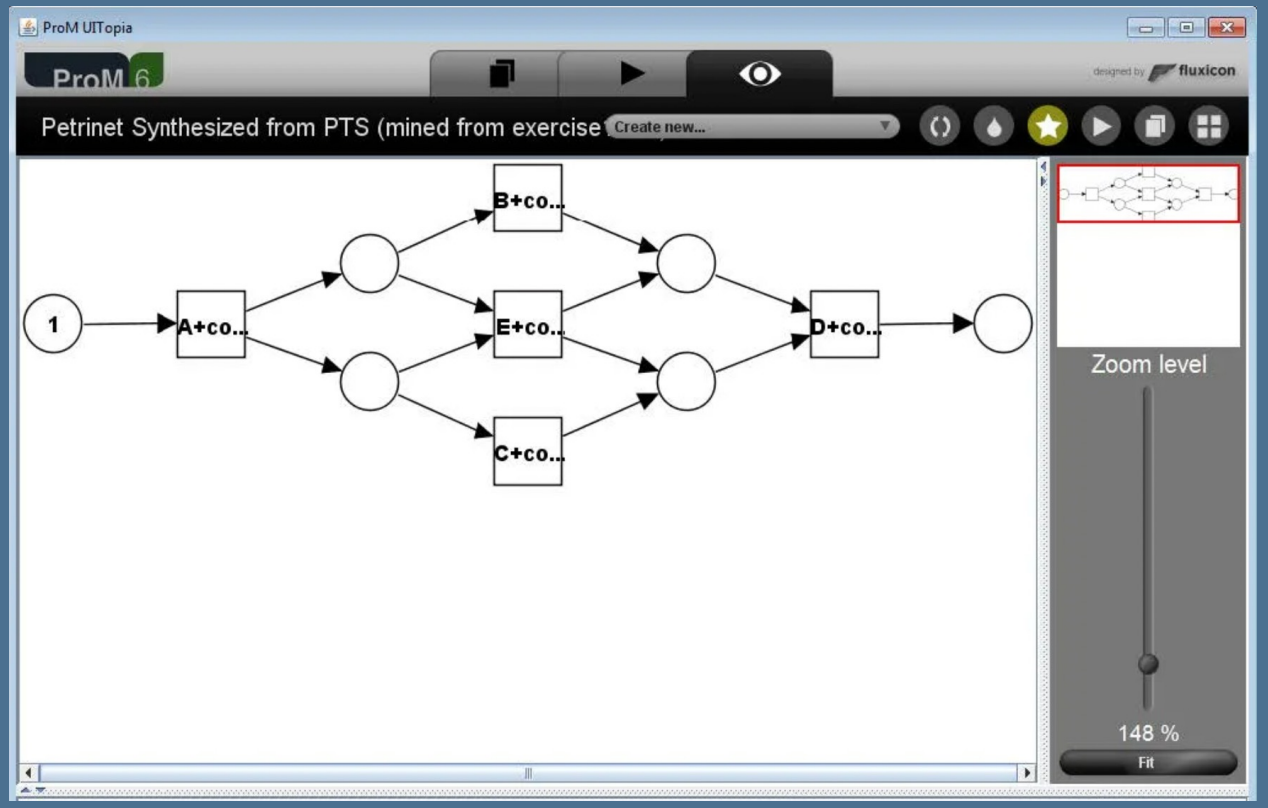


Fig. 9 The interface of ProM application

Overview of existing solutions

- PMTK (novel web-based Process Mining ToolKit)
- Celonis (commercial online platform for analysis business processes)
- SankeyMATIC (website for building Sankey diagrams)
- Fluxicon Disco (application for analysis business processes)
- ProM (plugin-based framework for process mining)
- **Proceset (analytical system for collecting and uploading data, conducting research in process mining)**

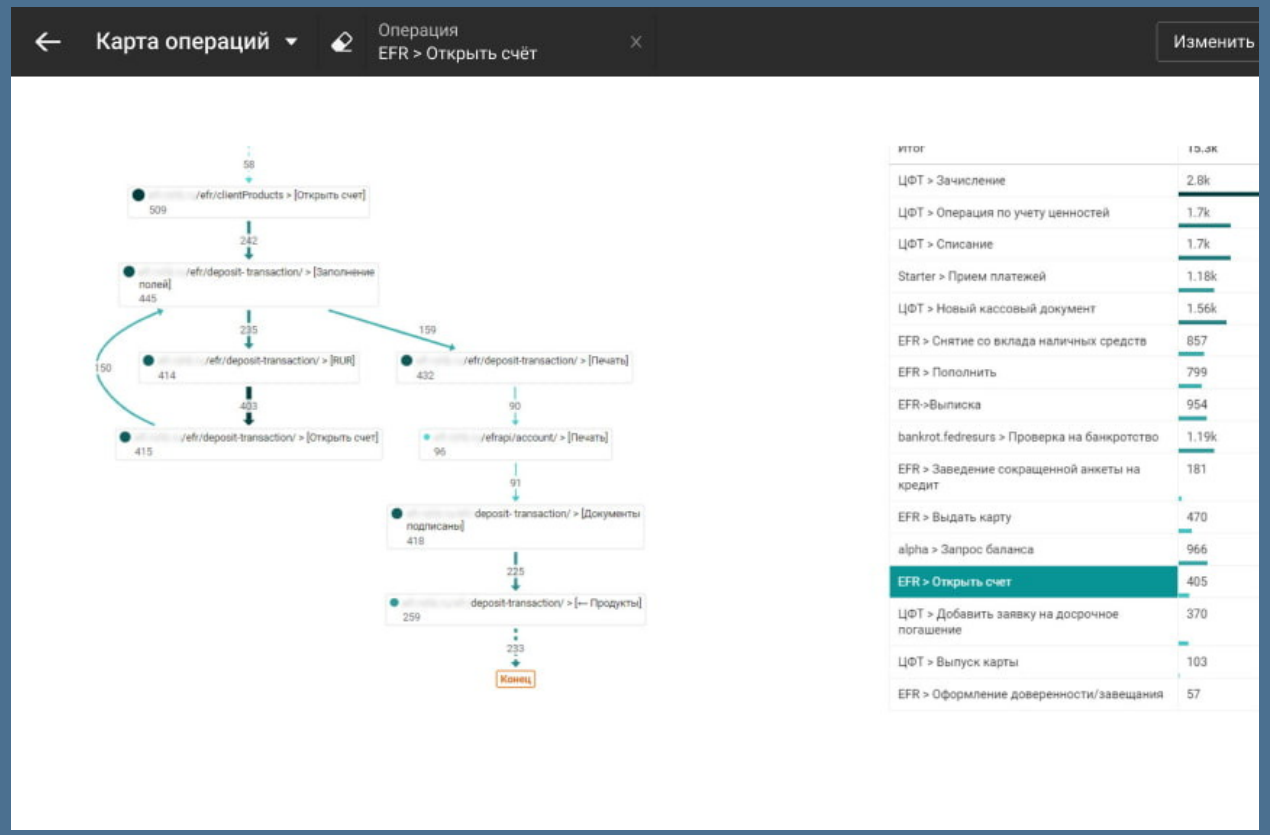


Fig. 10 The interface of Proceset application

Existing problems

1. The inability to filter the event log (graph) by trajectory options through interactive selection of events and transitions (by clicking on the vertices and edges of the graph);
2. The inability to create and save subsets of cases;
3. **The inability to visualize for models with intersecting subsets;**
4. **The inability to visualize multiple subsets on the same graph.**

Motivating example

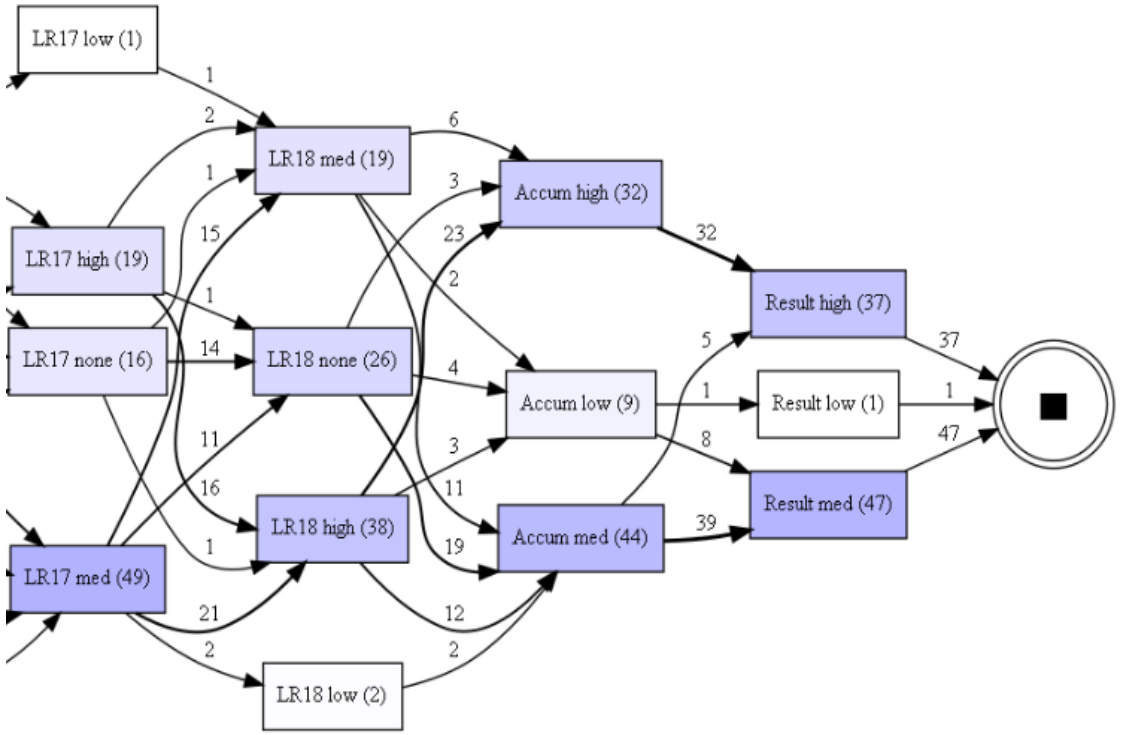


Fig. 11 The resulting DFG model

Case_id	Activity	Timestamp	Grade	Lector_id	Group
#19273	LR1 high	08.11.2021	8	#13761	BPI203
#19273	LR2 med	09.11.2021	6	#13761	BPI203
#19273	LR3 med	16.11.2021	5	#13761	BPI203
…					
#45951	LR17 med	11.12.2021	6	#13976	BPI206
#45951	LR18 low	17.12.2021	3	#13976	BPI206

Table 1 The example of an eventlog

Motivating example

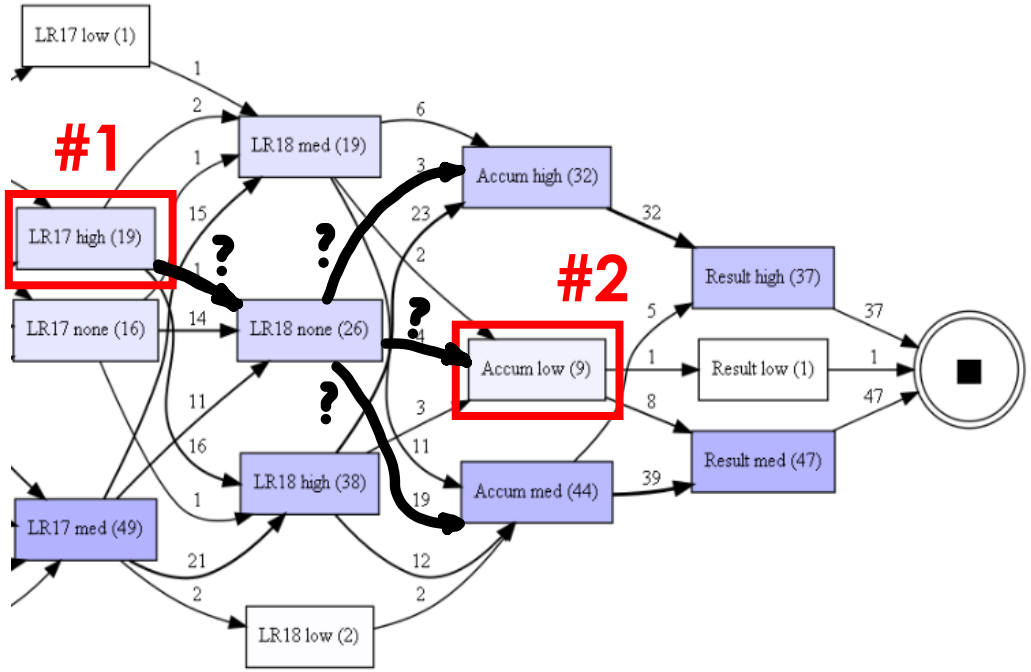


Fig. 12 The resulting DFG model

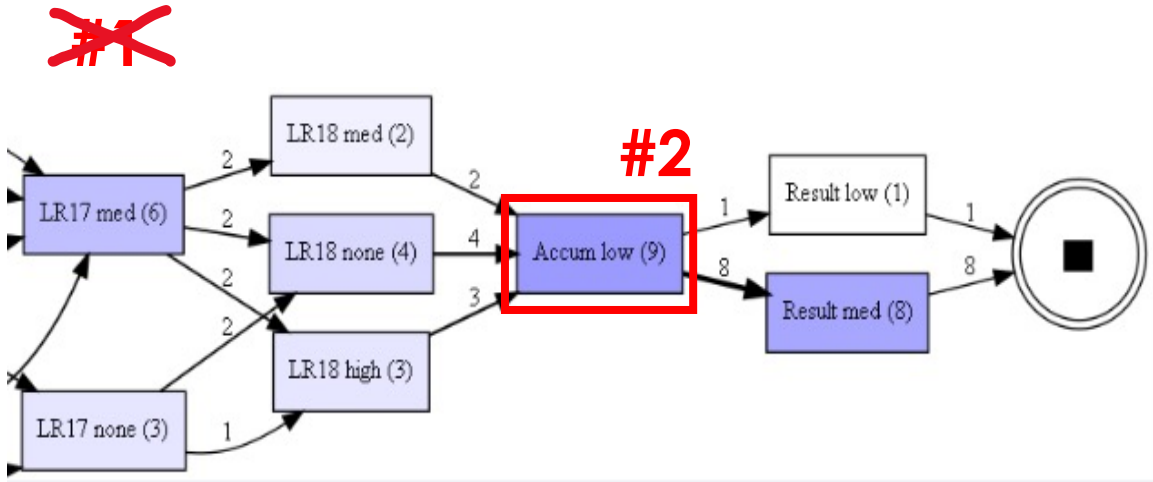


Fig. 13 The DFG model for traces containing the «Accum low» event

Proposed solution

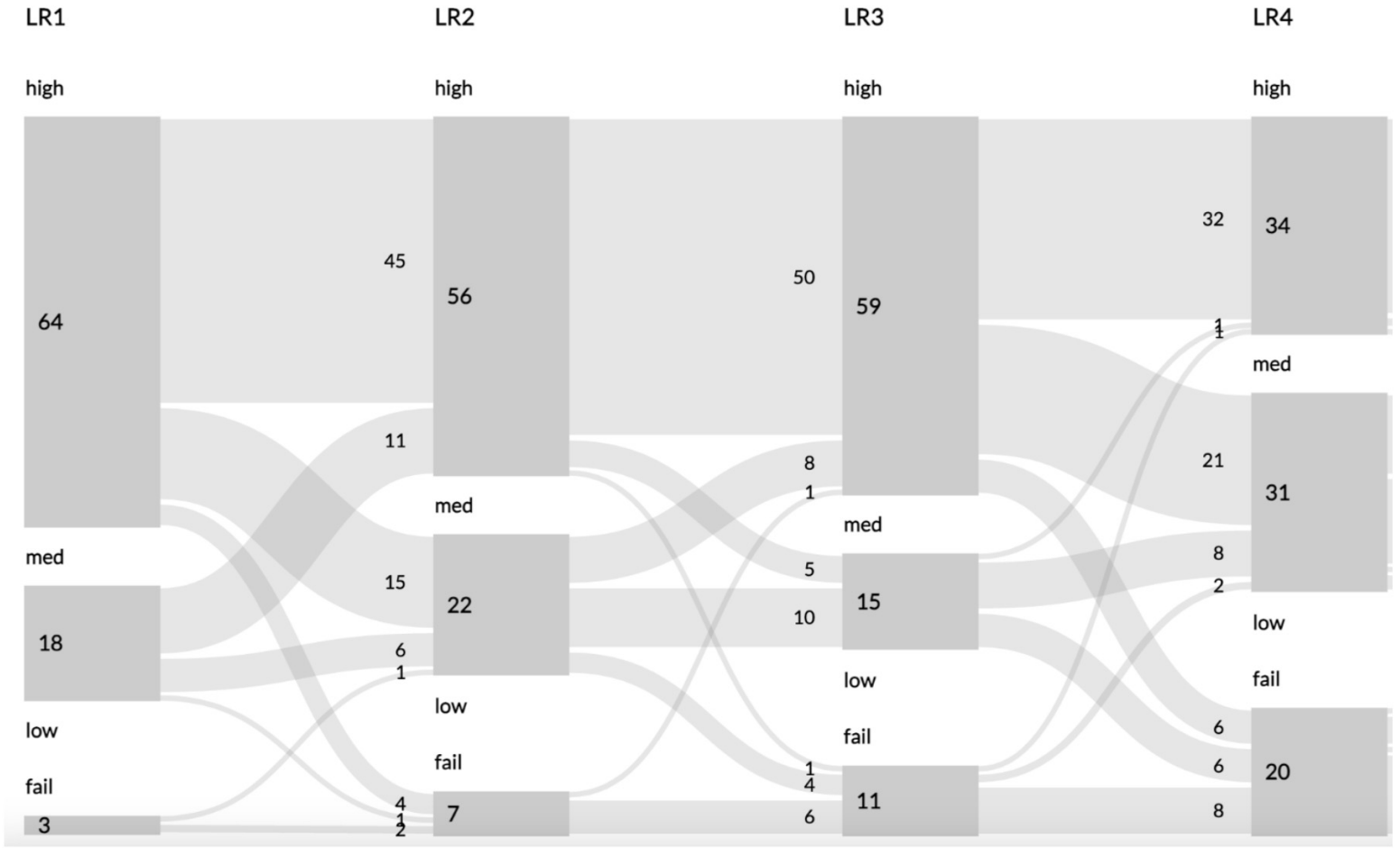


Fig. 14 The DFG model merged with the Sankey diagram

DFG
+
Sankey diagram

Visualization of the flow direction

The width of the transition lines reflects the size of the transition

The ability to display multiple streams

Convenience in analyzing the structure and effectiveness

Proposed solution

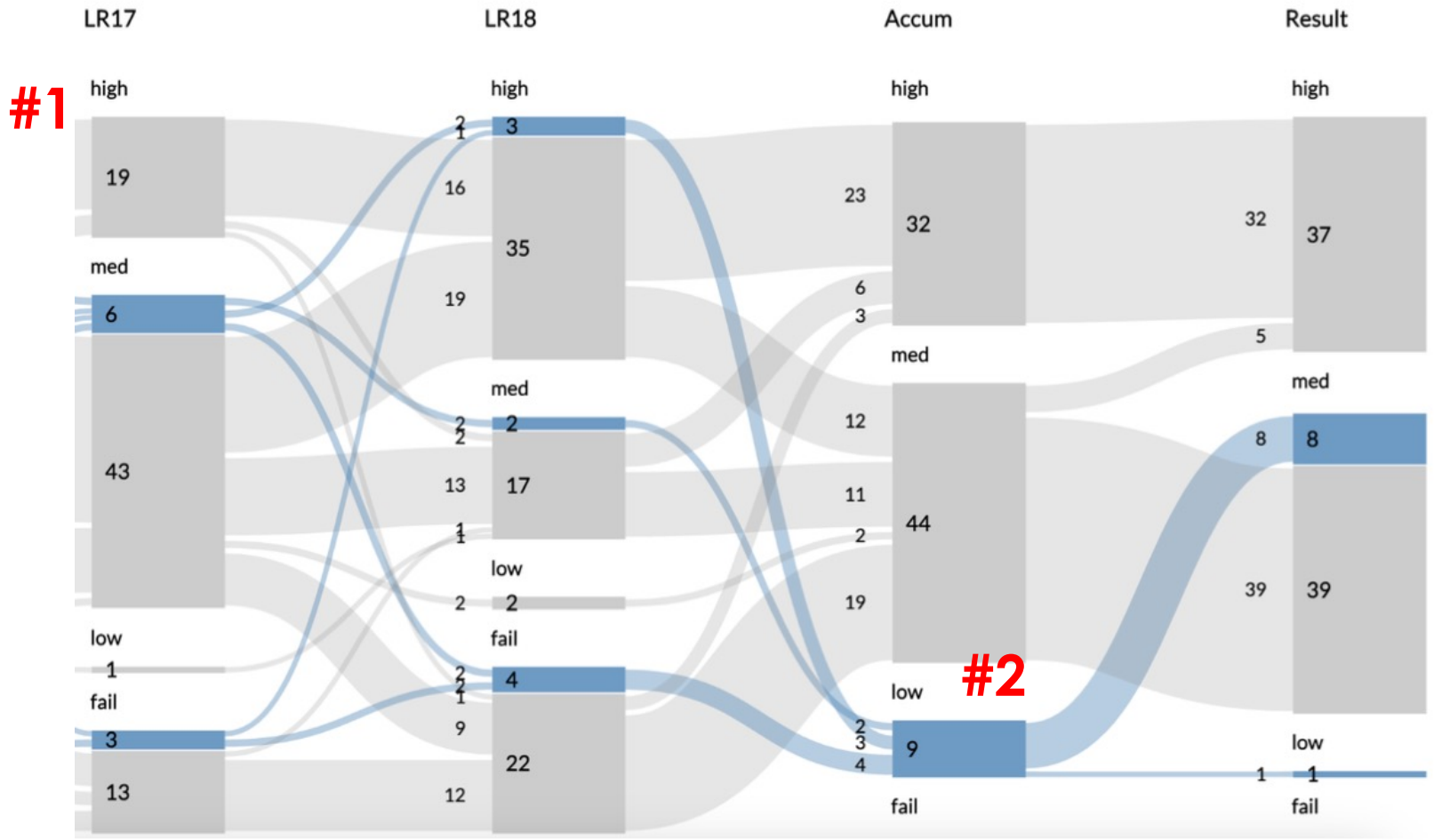


Fig. 15 The DFG model merged with the Sankey diagram with traces containing the «Accum low» event highlighted

Proposed solution

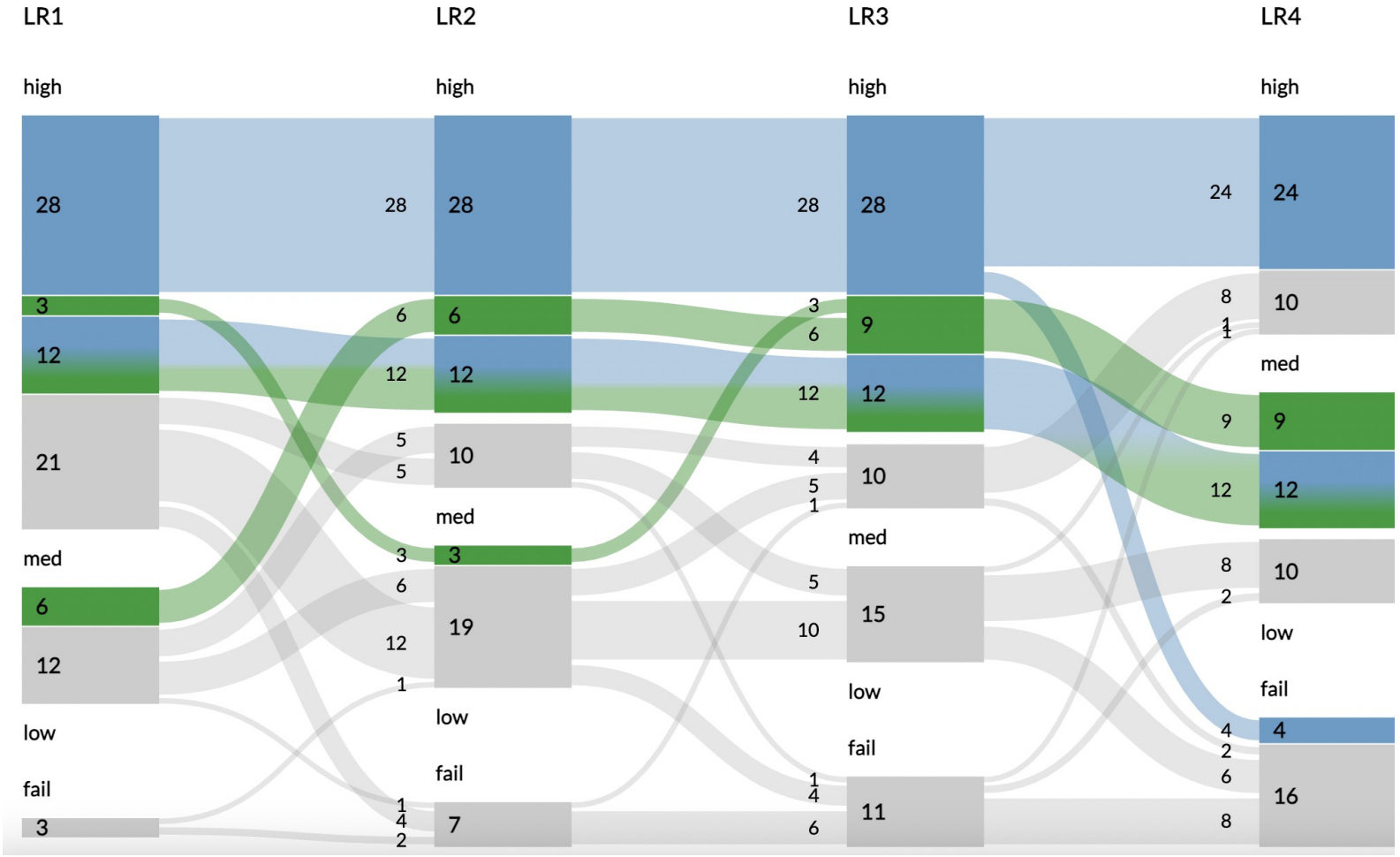


Fig. 16 Visualization of the intersection of two subsets of cases

Visualization can display 4 different categories of cases that:

- present only in the first subset (**blue** color);
- present only in the second subset (**green** color);
- present in both subsets (**gradient** color);
- not present in any of the subsets (**grey** color).

Tool architecture and implementation

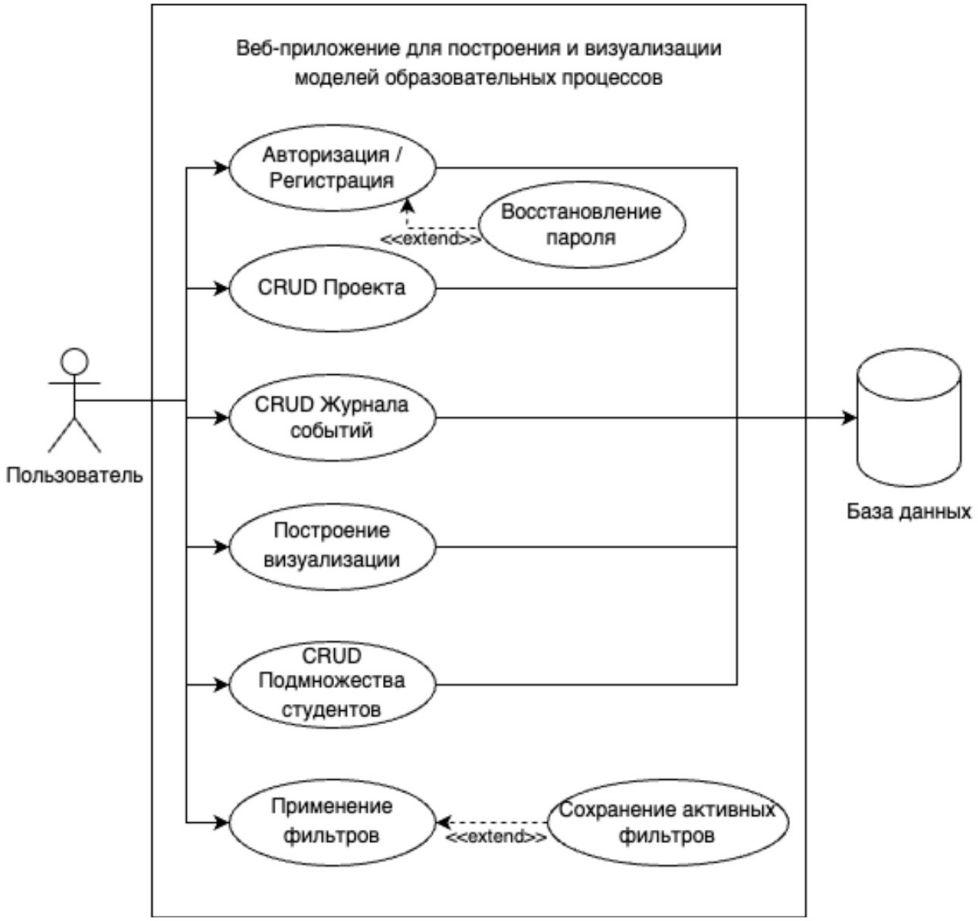


Fig. 17 The Use case diagram

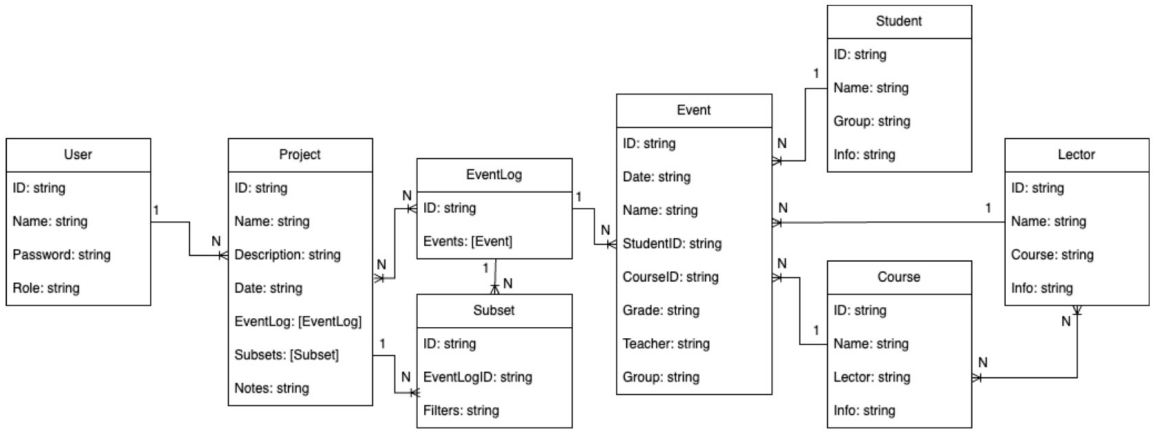
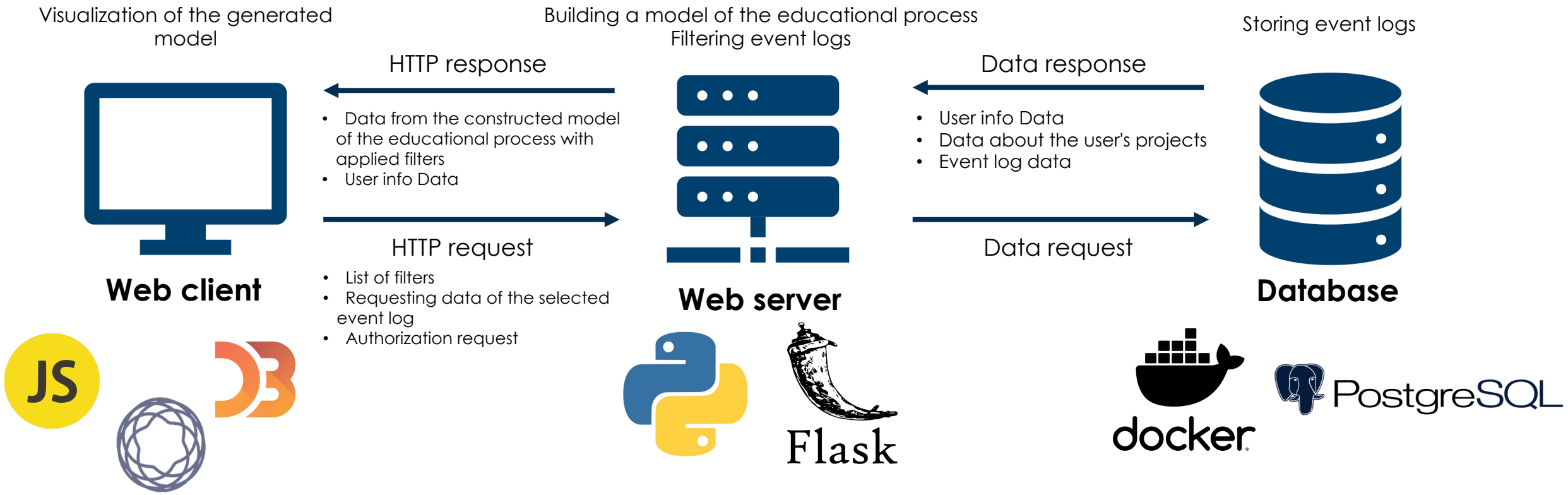


Fig. 18 The domain model of the system

Tool architecture and implementation



Findings and Evaluations

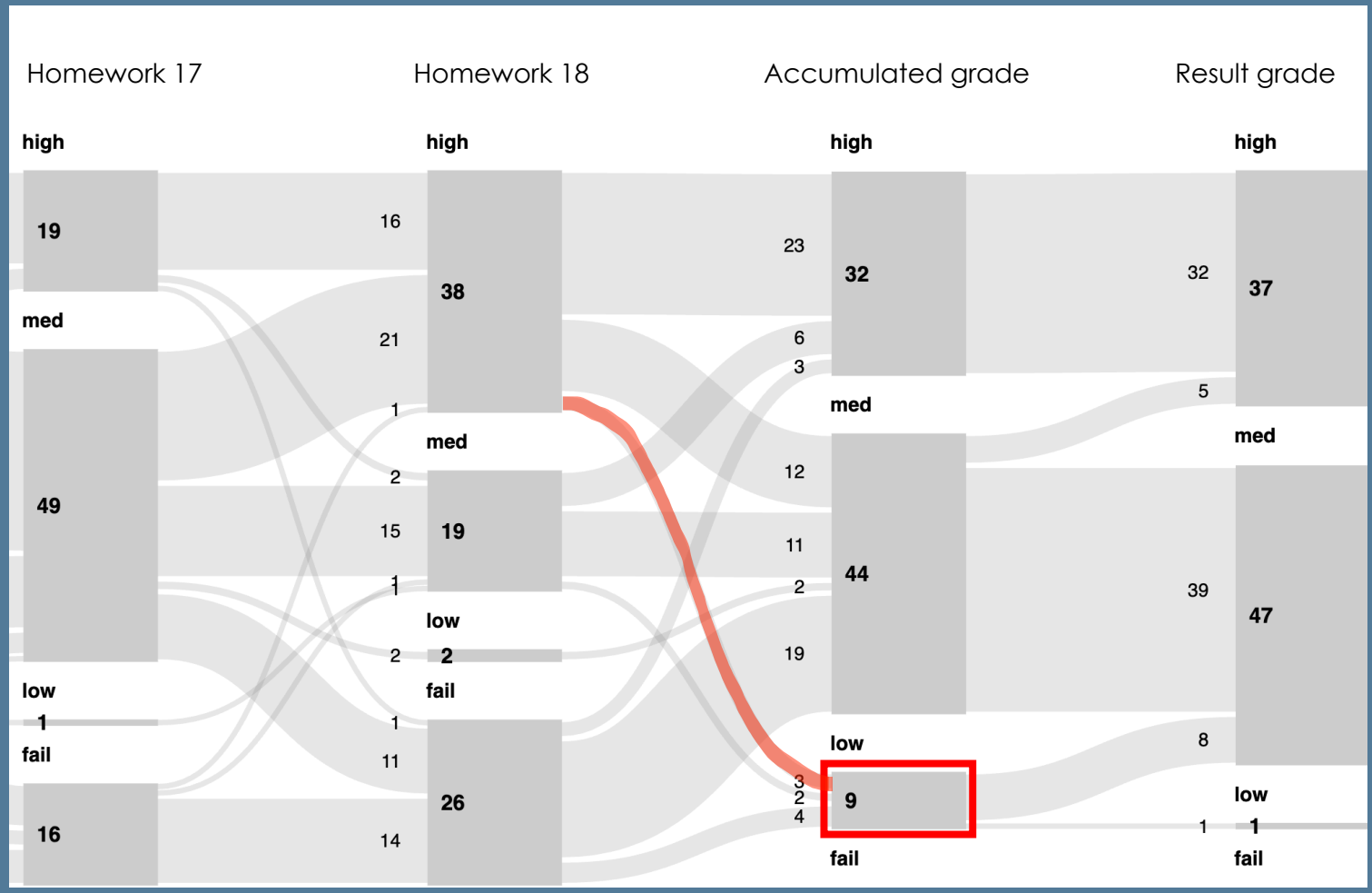


Fig. 19 The resulting DFG model merged with the Sankey diagram

Findings and Evaluations

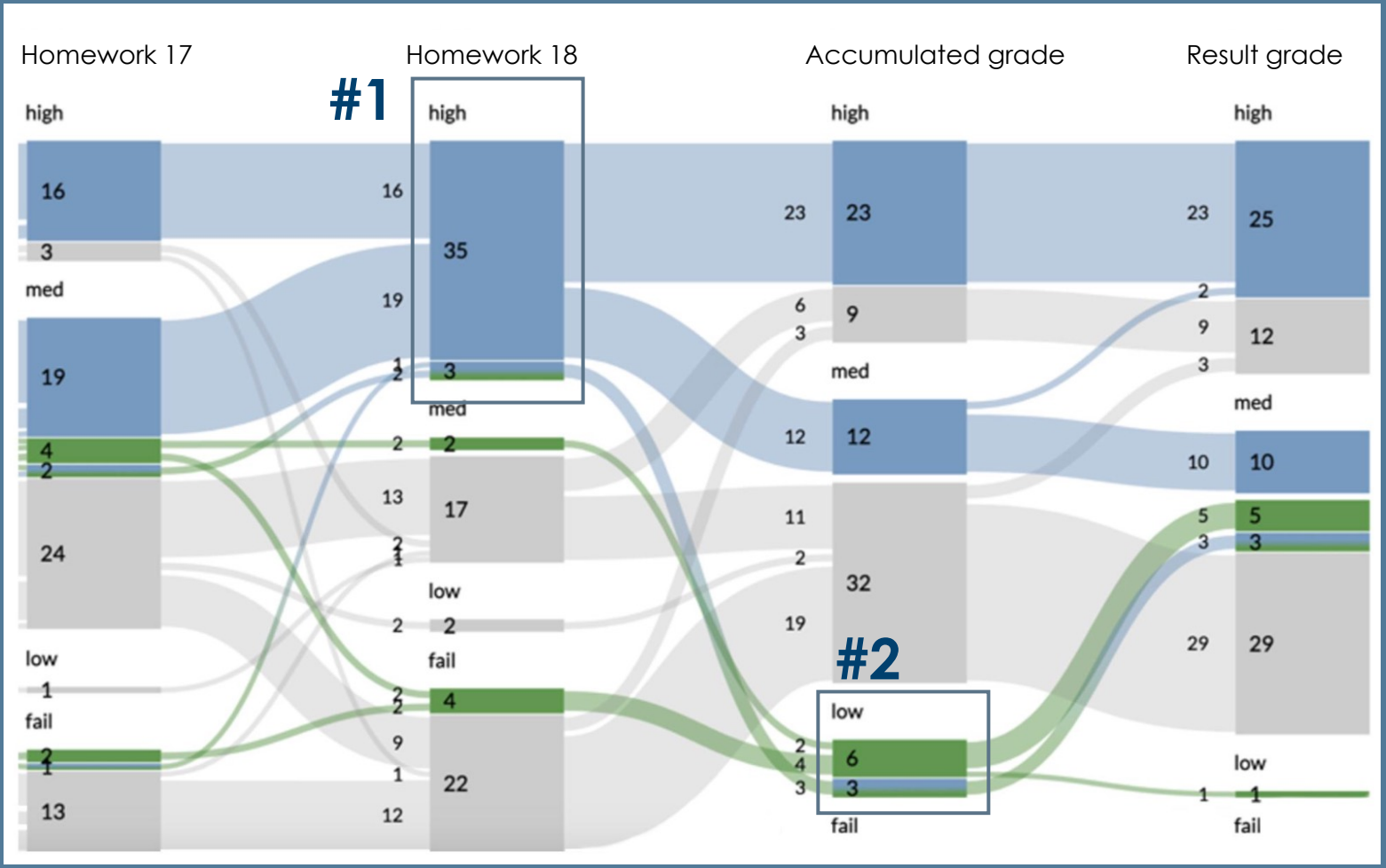


Fig. 20 Visualization of the intersection of two subsets of cases

- First subset (**blue** color) - students who received an excellent grade for the last work (node «high Homework 18»);
- Second subset (**green** color) - students who received a low accumulated grade (node «low Accumulated grade»).

Findings and Evaluations

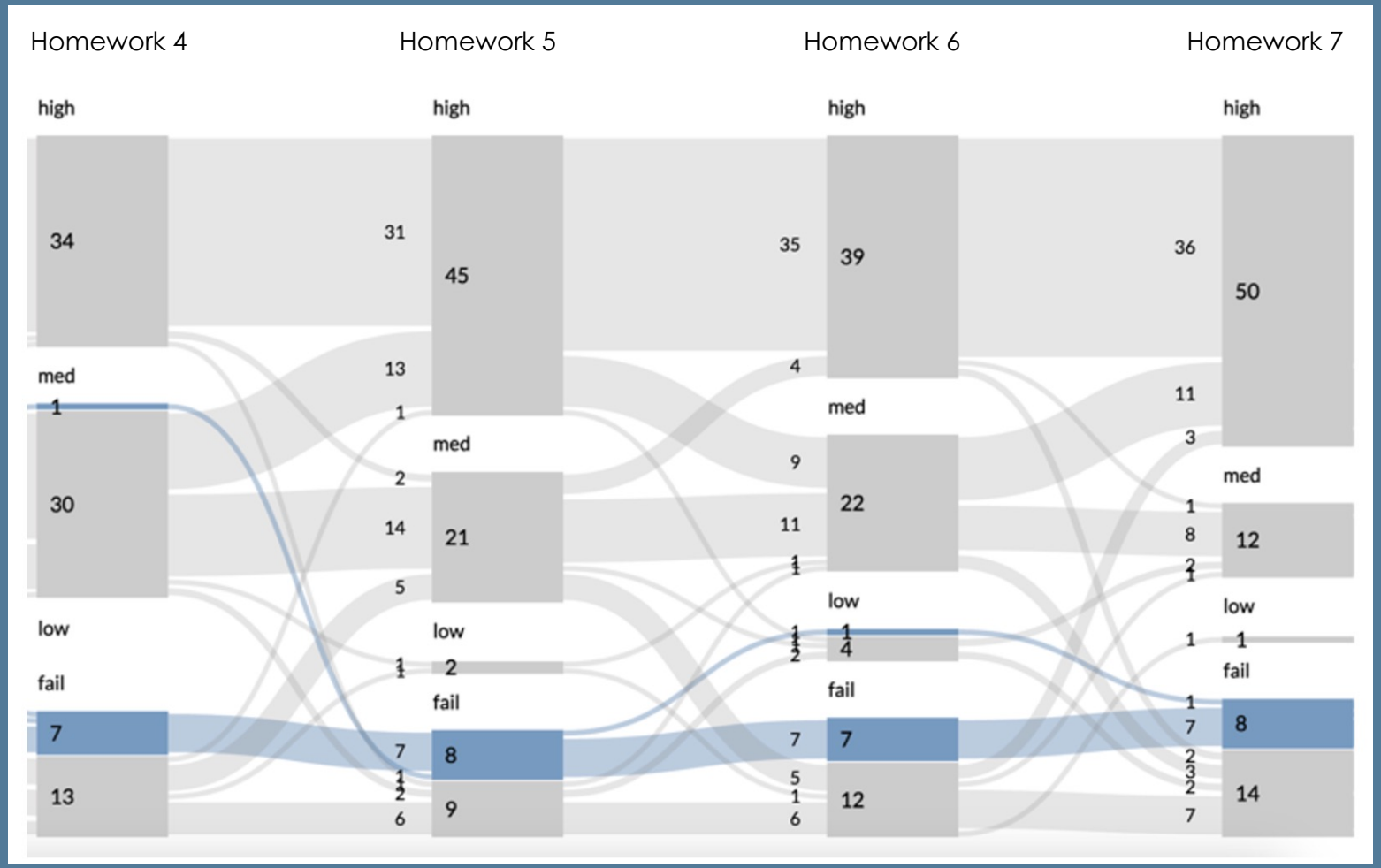


Fig. 21 Visualization of the subset of students

Future plans

- Generalization of our visualization method for models with cycles
- Adding new functionality for filtering and visualizing data
- Testing of the application by employees and managers of educational programs of the FCS HSE

Merging Directly-Follows Graphs and Sankey Diagrams for Visualizing Acyclic Processes

Ilya D. Derezovskiy, Nikita D. Shaimov, Irina A. Lomazova, Alexey A. Mitsyuk

HSE University, Faculty of Computer Science, PAIS Lab
Moscow, Russia

iderezovskiy@hse.ru, nshaimov@hse.ru, ilomazova@hse.ru, amitsyuk@hse.ru

SYRCoSE 2024
May 30, 2024