

Kann uns AI helfen, besser zu testen?

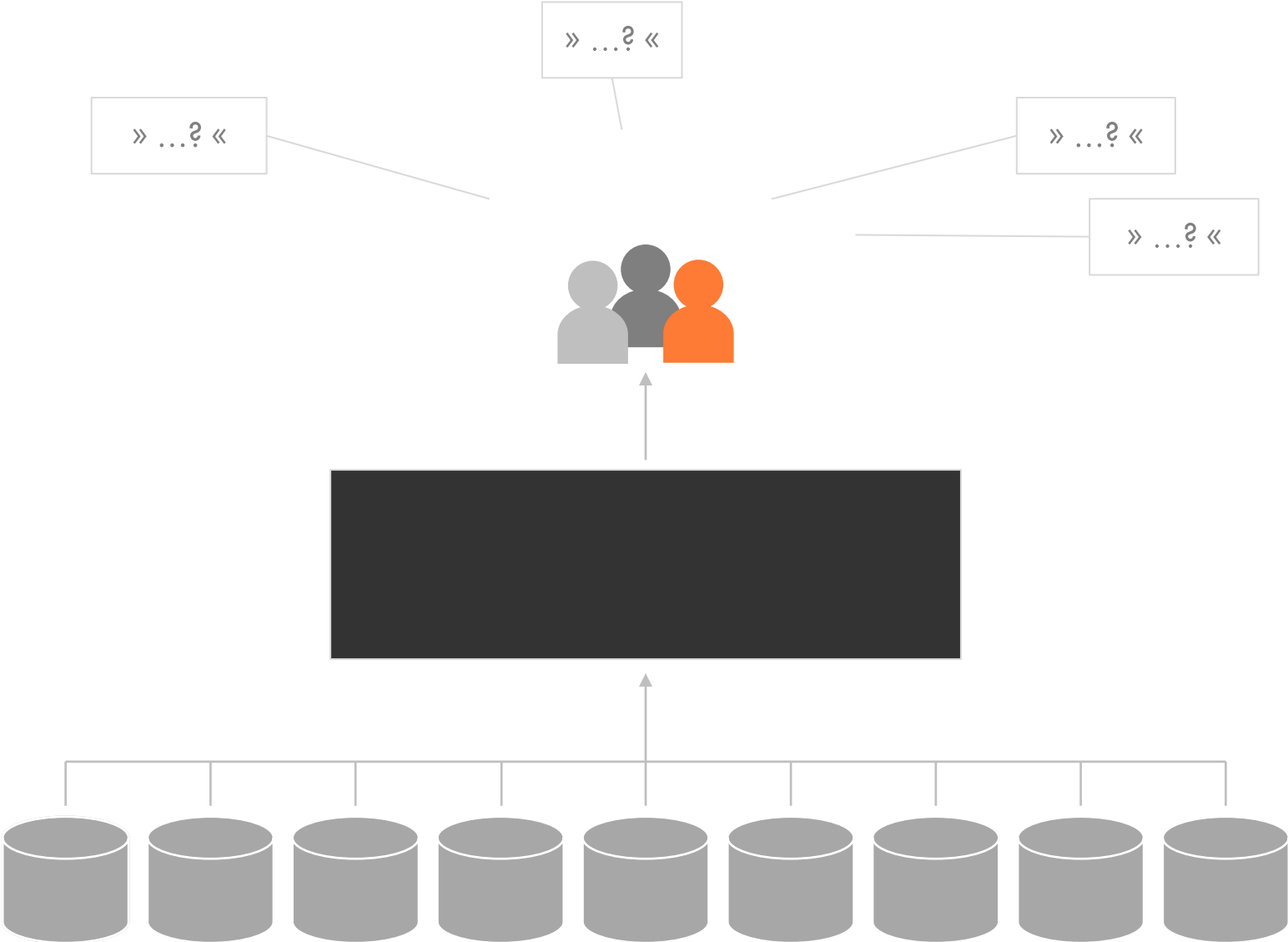
Erfahrungen mit *Machine Learning* in der QS in Forschung und Praxis



TUM



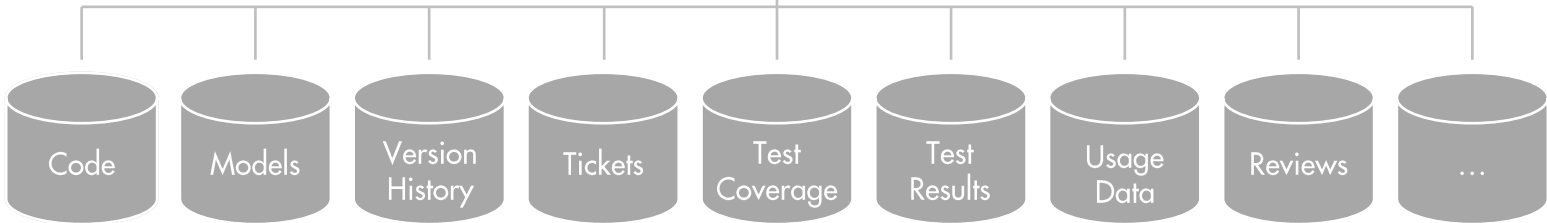
CQSE



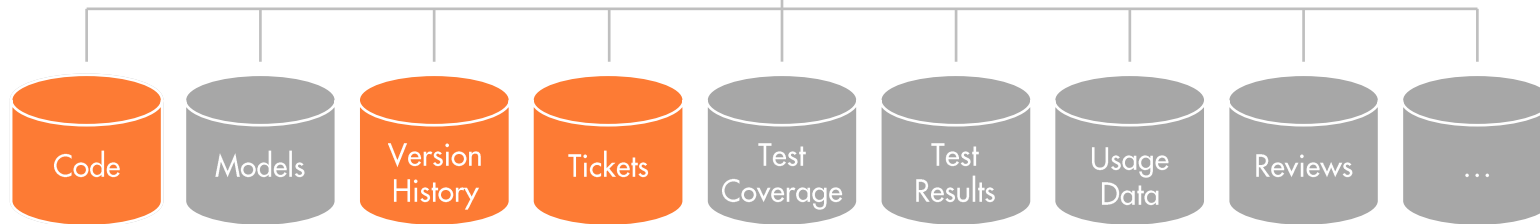
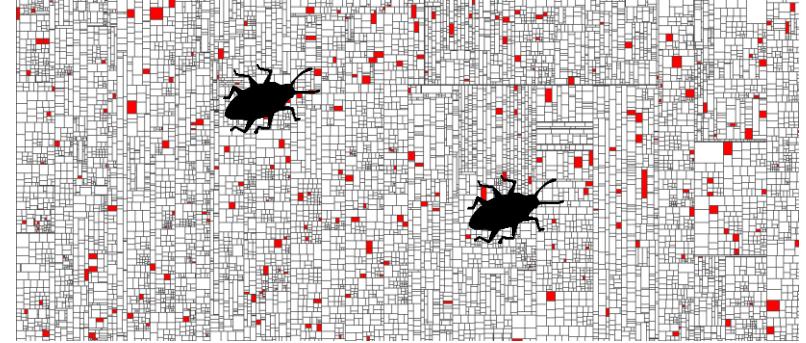
» Wo sind Fehler im Code? «

» Welche Tests sollte ich jetzt ausführen? «

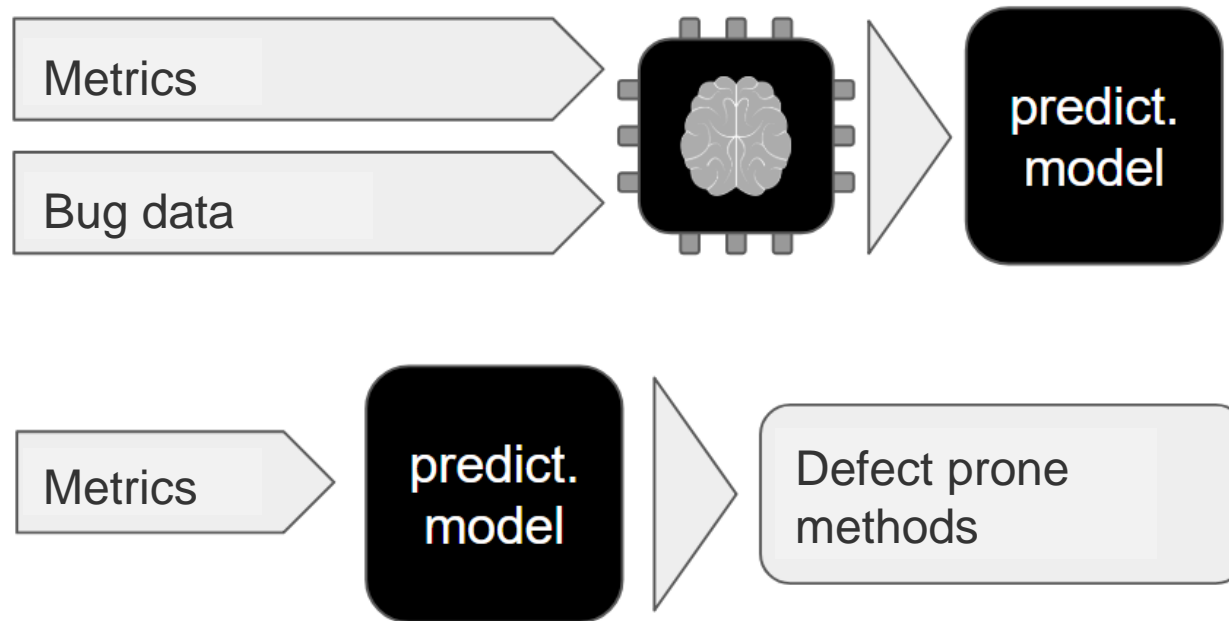
» Wo sind ungetestete Änderungen «

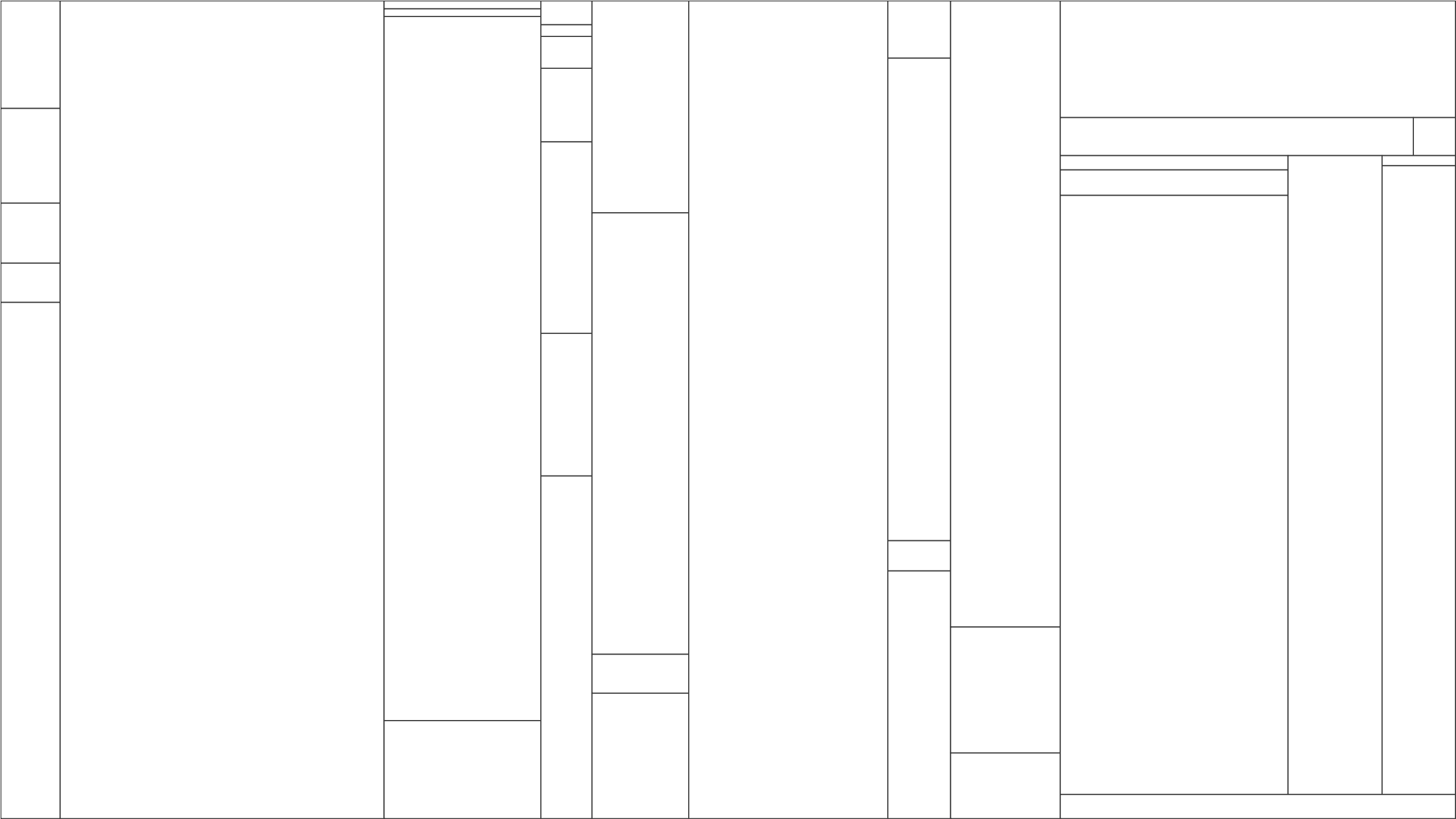


» Wo sind Fehler im Code? «

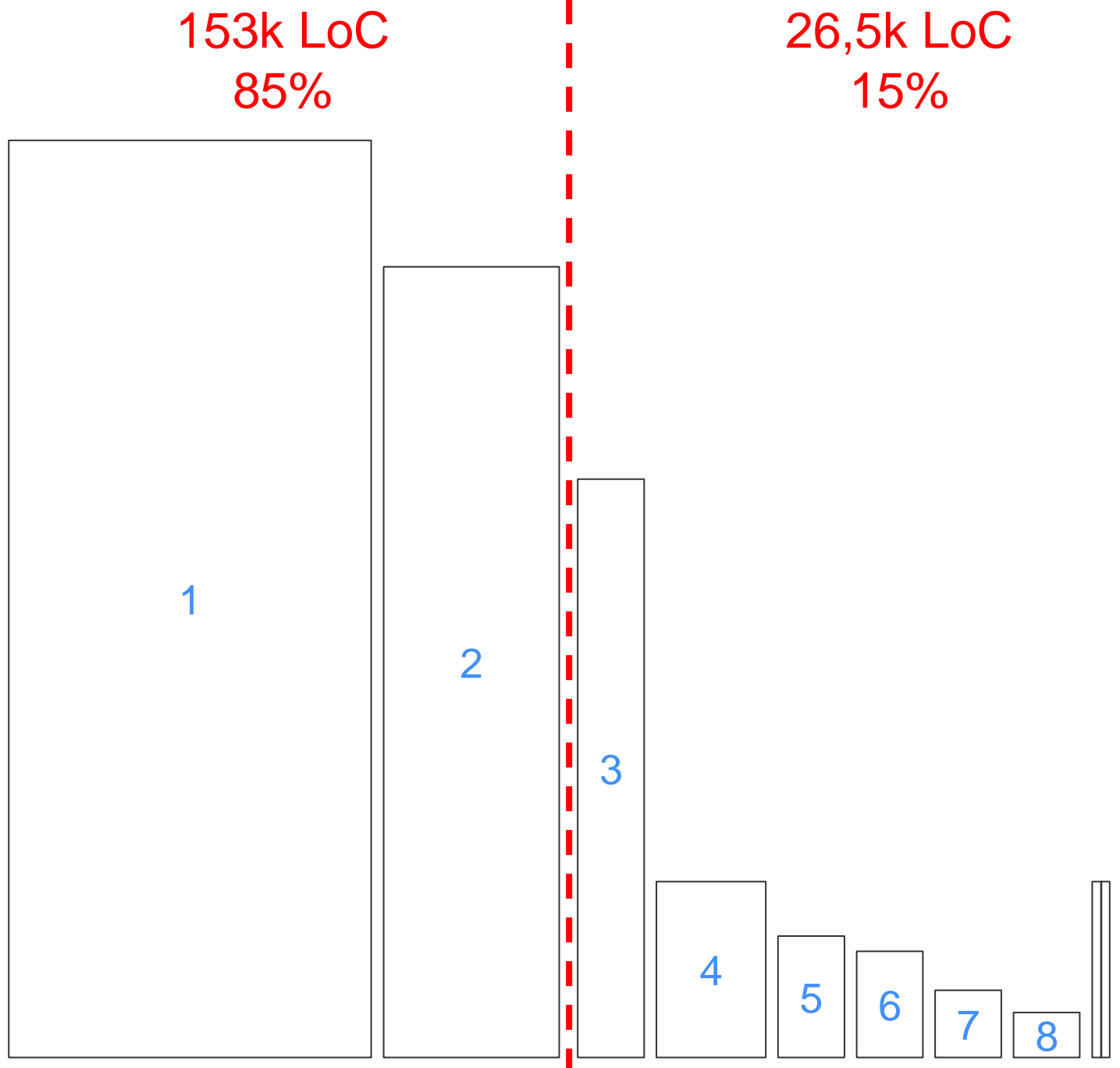
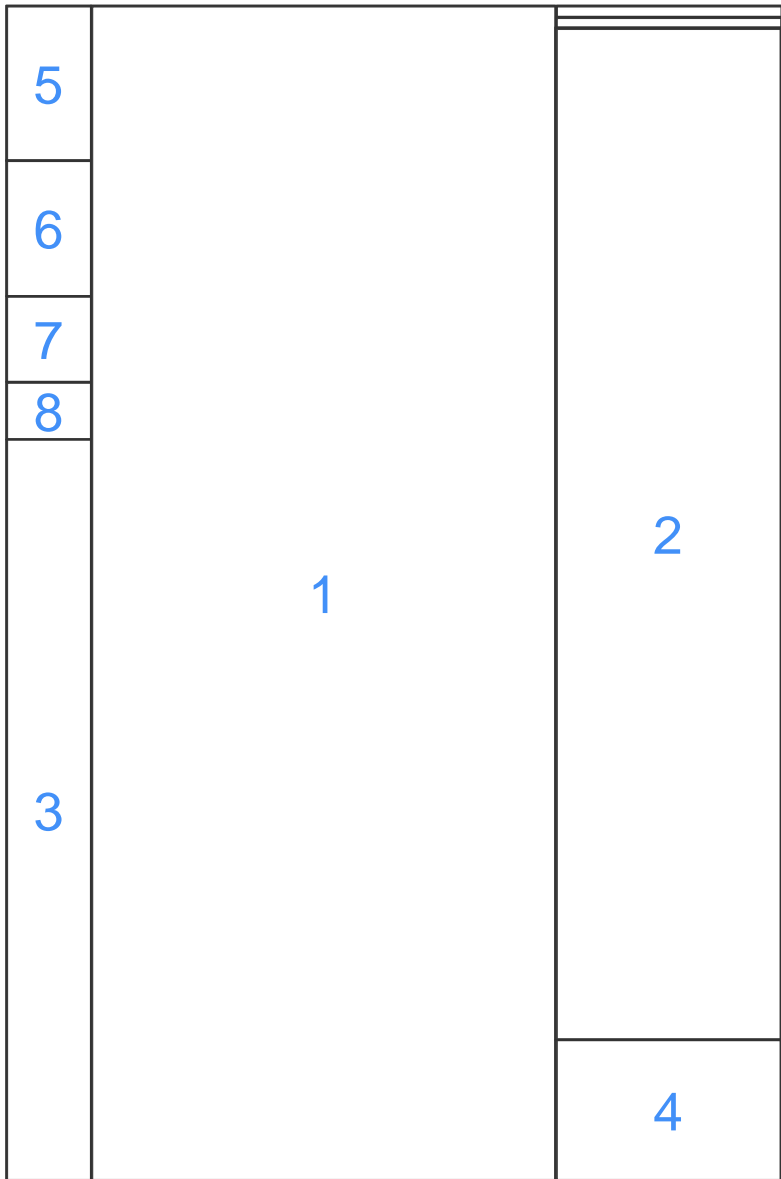


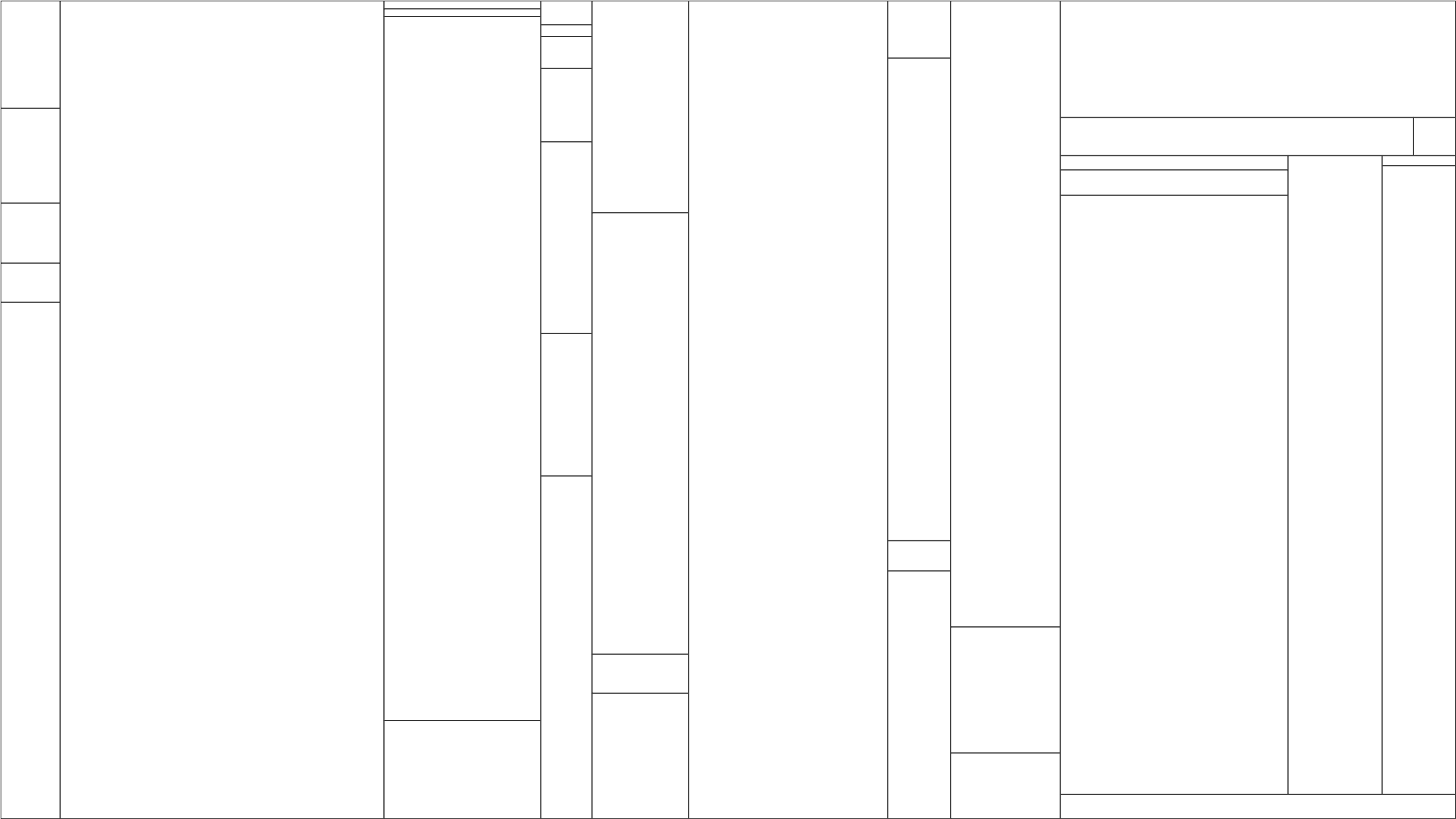
Defect Prediction

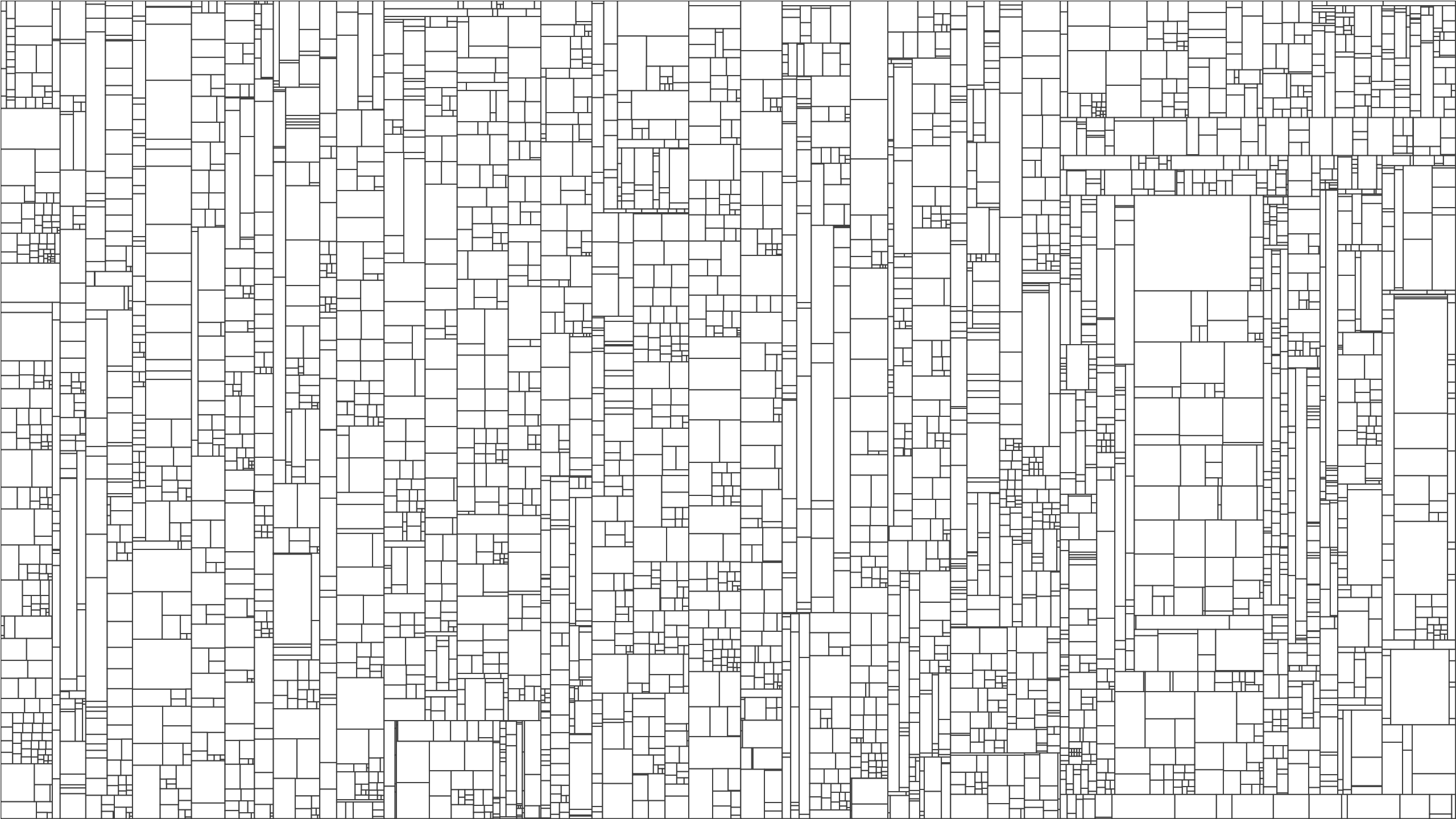


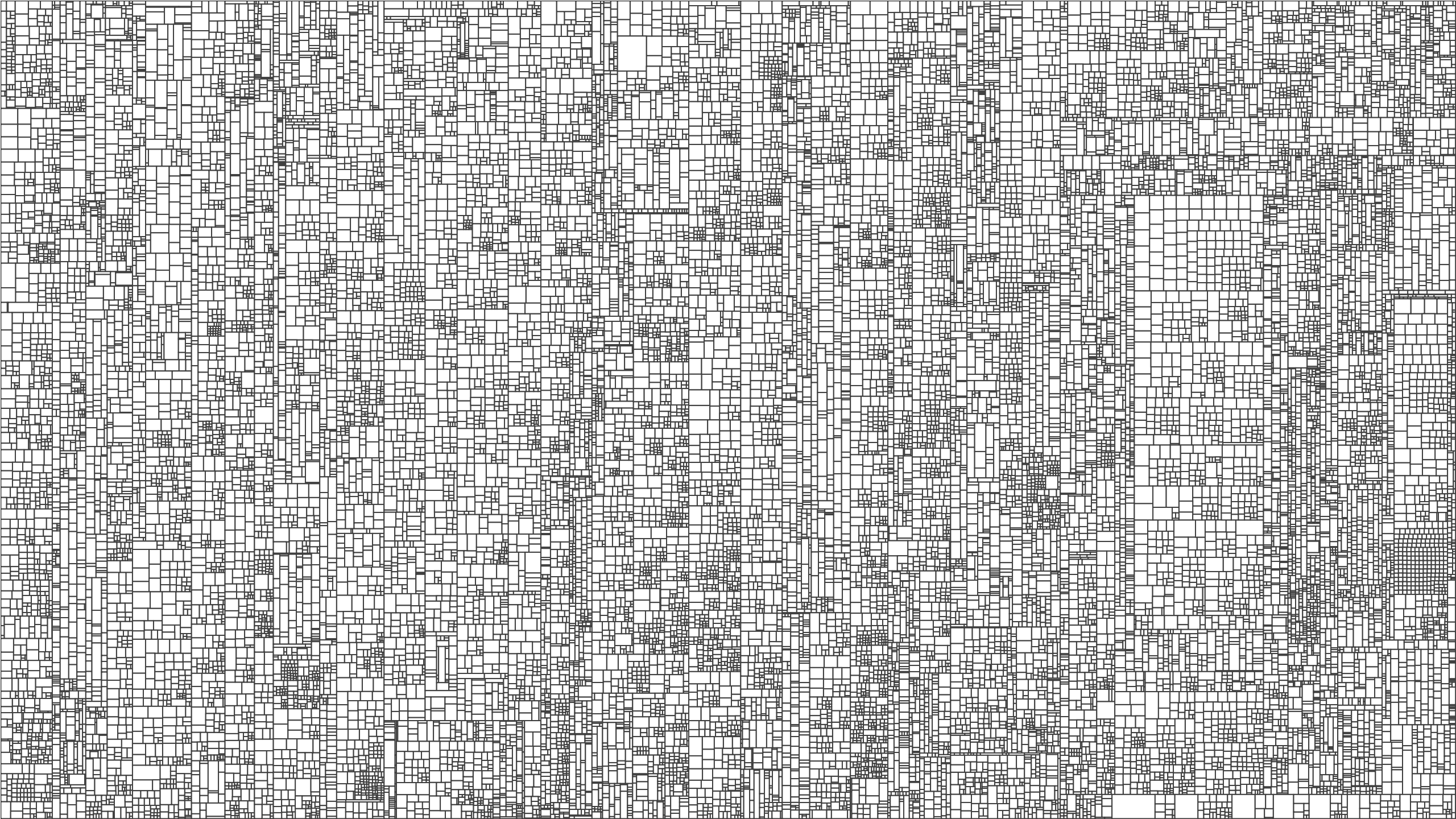


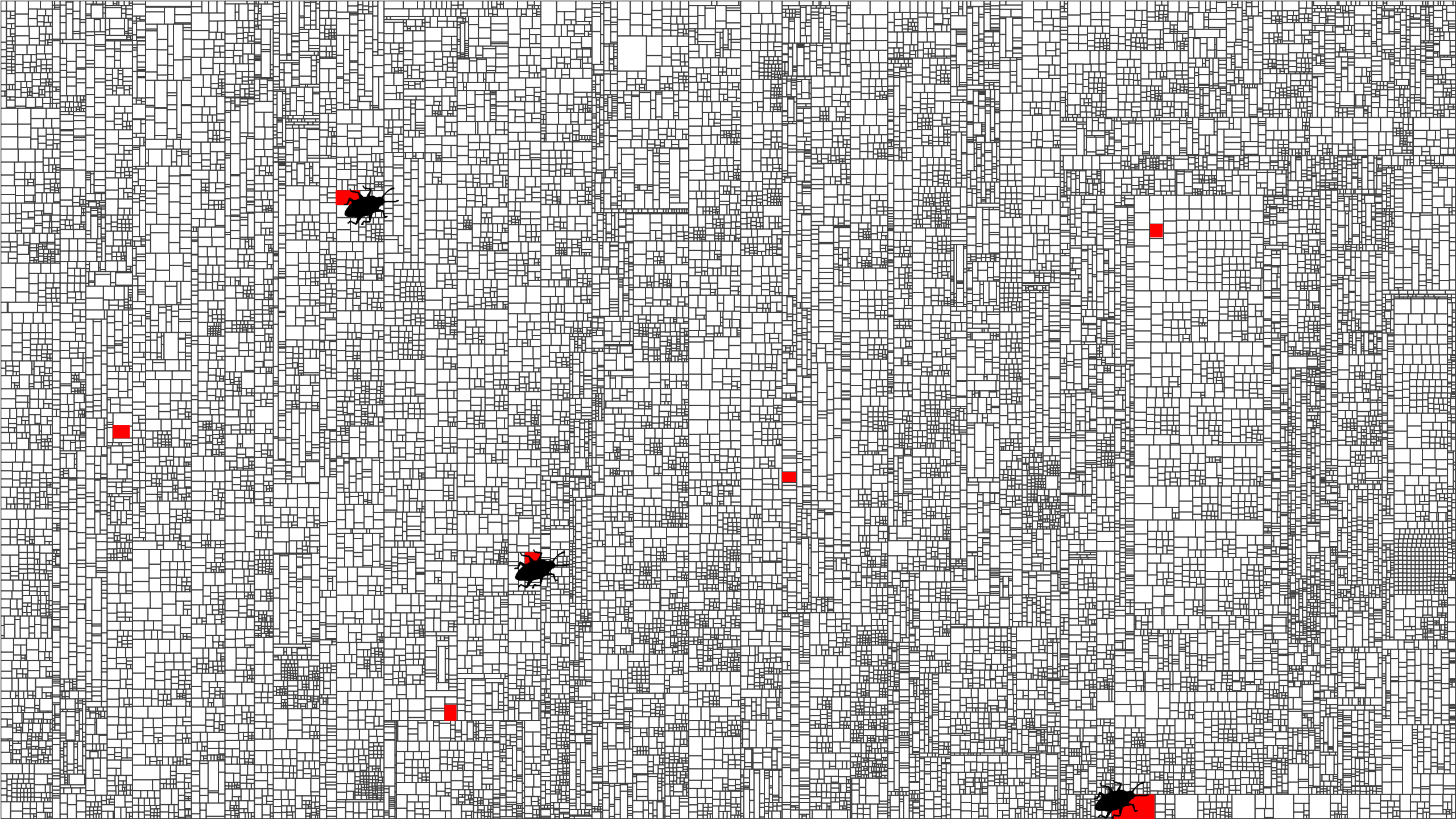
179k LoC

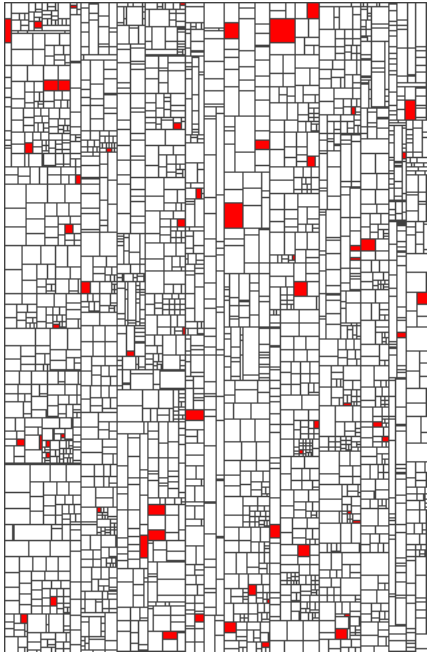






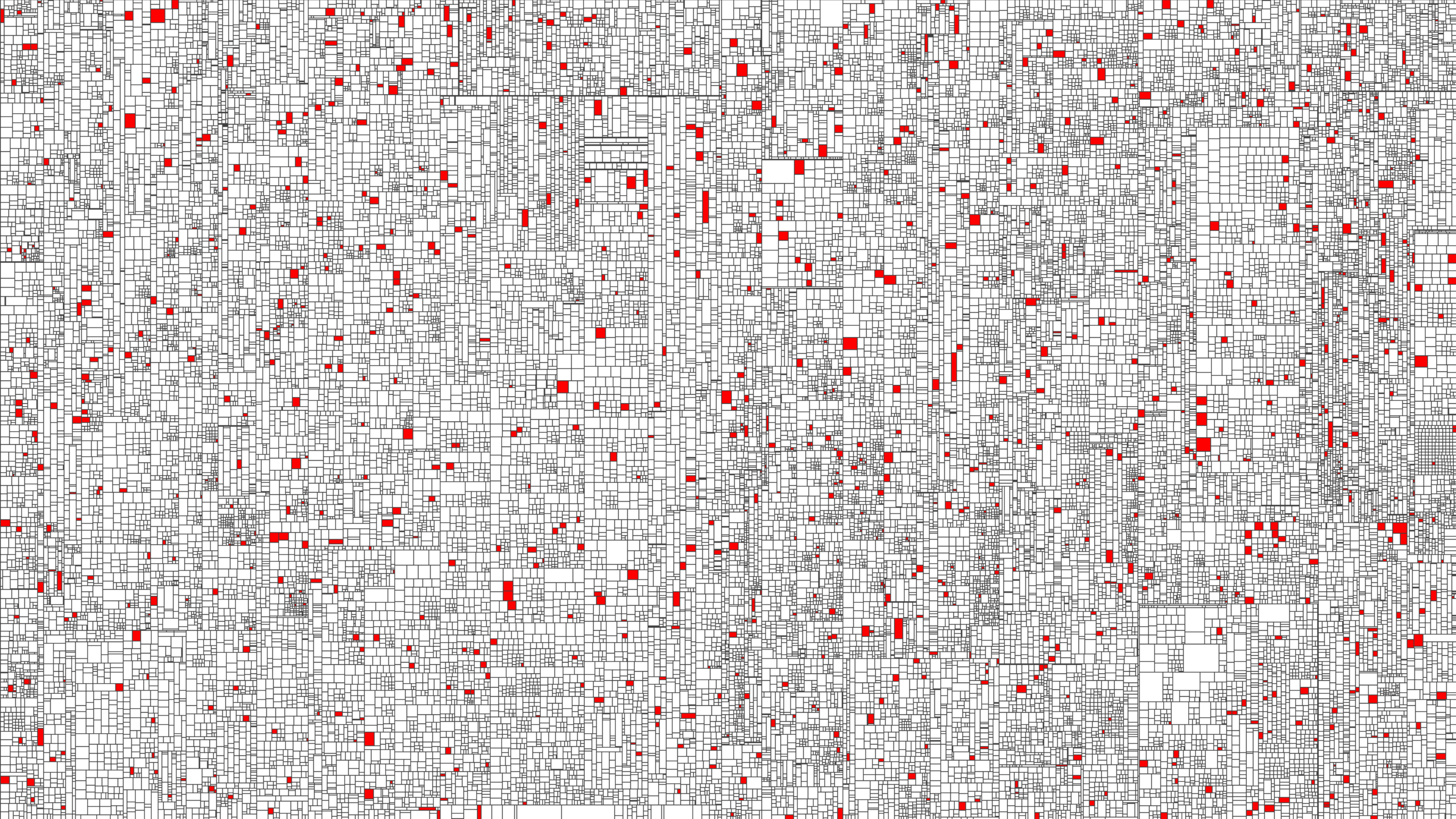


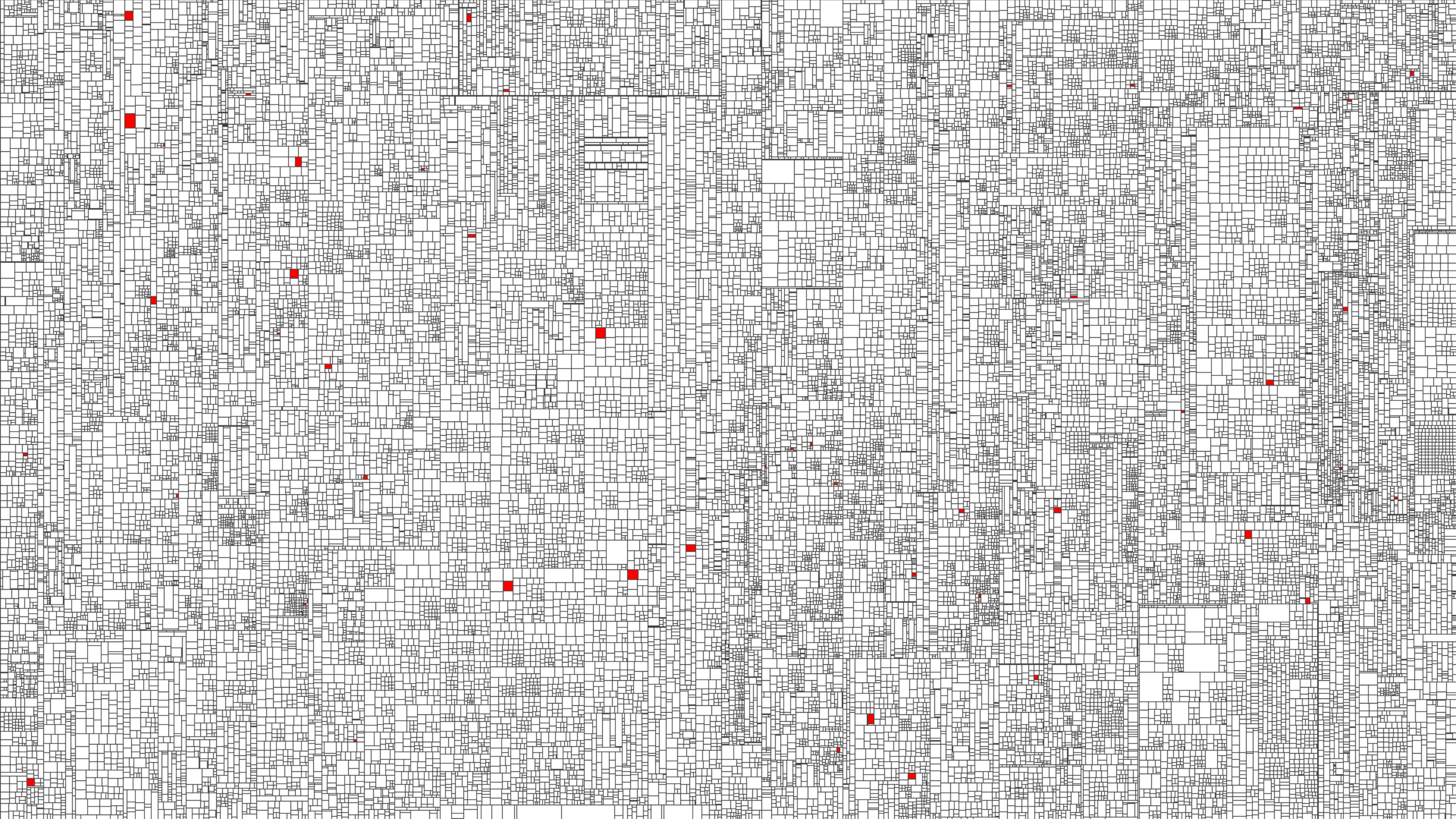




Release

Time of Study

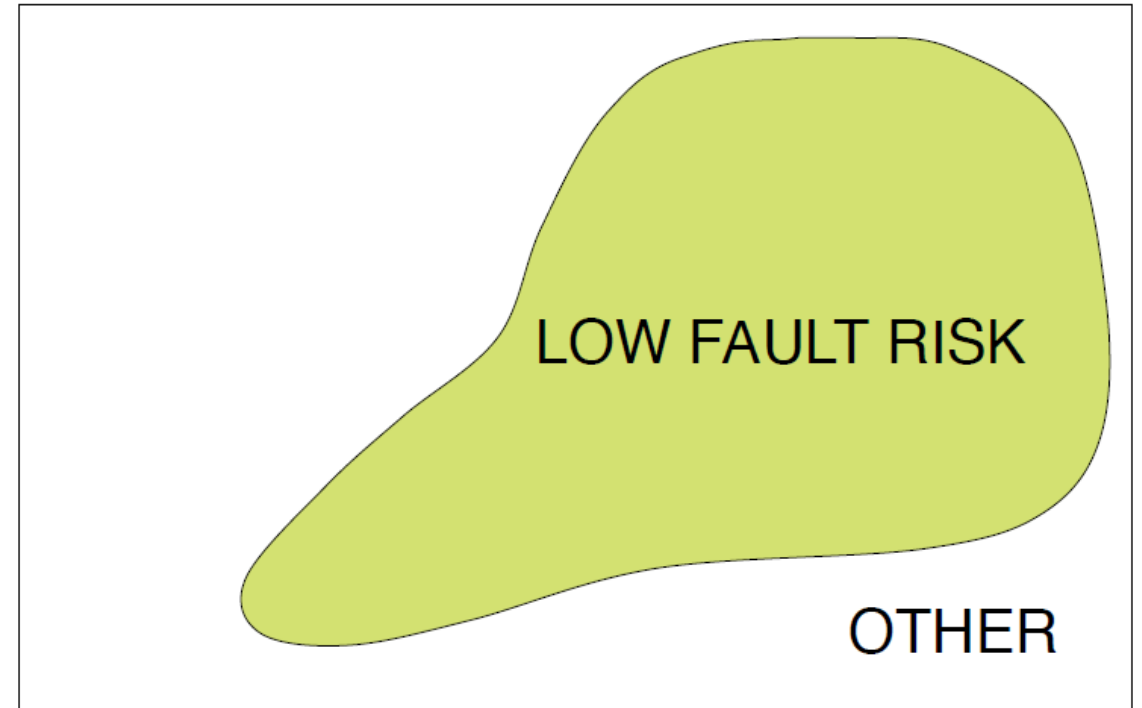
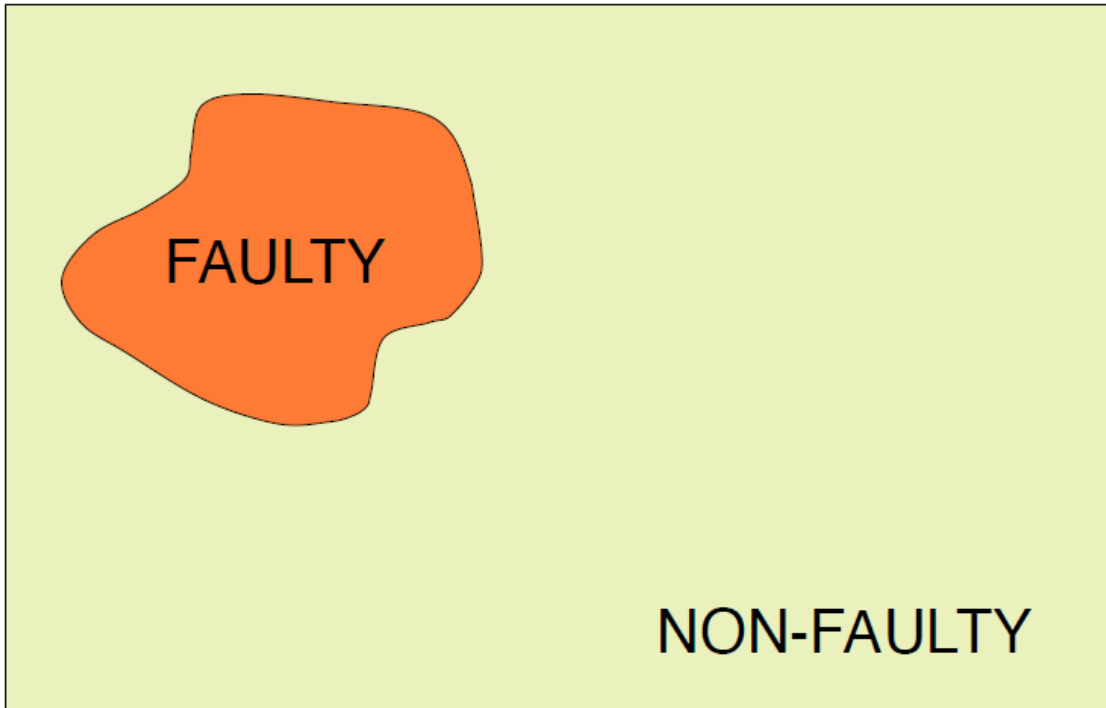




Evaluation

Release	# „defect prone“ Methods	# Bugs (Top 50)
1.4:	1127	0
2.0:	1176	0

Pascarella, Palomba, Bacchelli, *Re-evaluating Method-Level Bug Prediction*, 2018:
Prediction nicht besser als zufällige Klassifikation.



Zentrale Ergebnisse

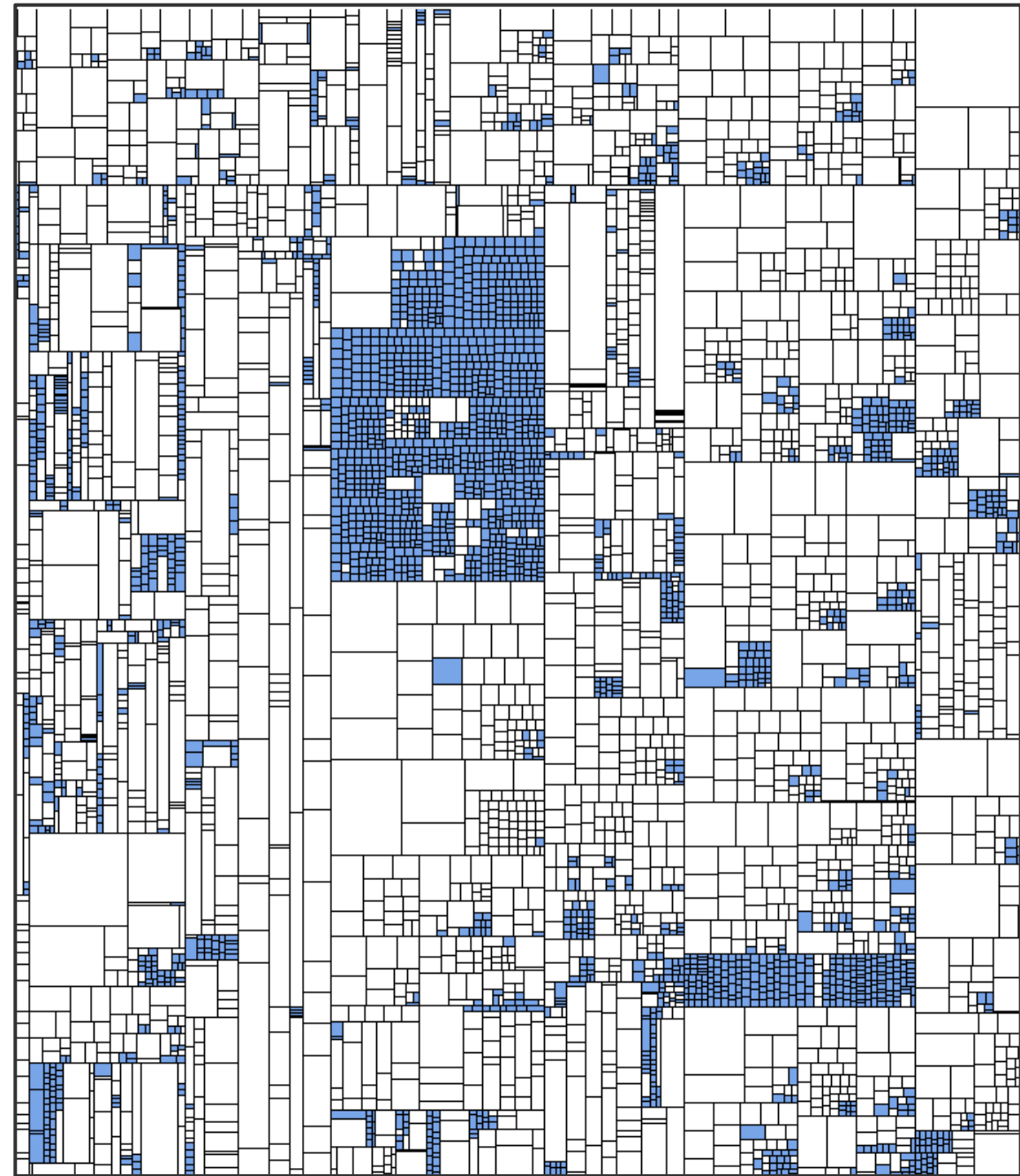
Low Fault Risk

- 16%-80% aller Methoden (Median: 31%)
- 5%-78% aller SLOC (Median: 15%)

Enthalten nur 6% der Fehler

Reduktion der Fehlerdichte

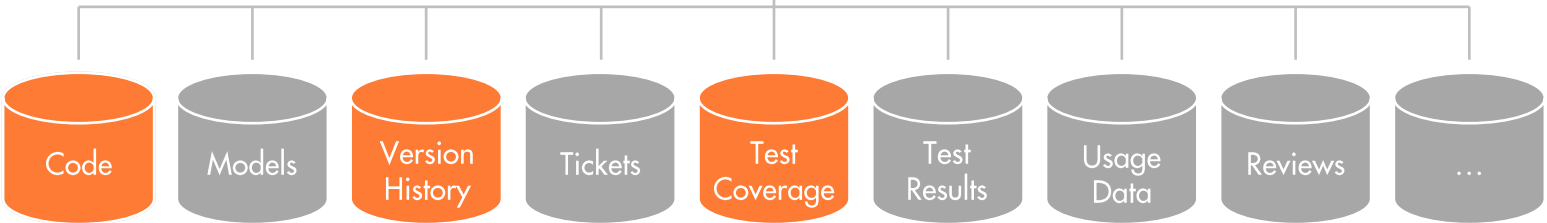
- 4,3 bis 10,9 fach pro Methode
- 1,5 bis 4,4-fach pro SLOC



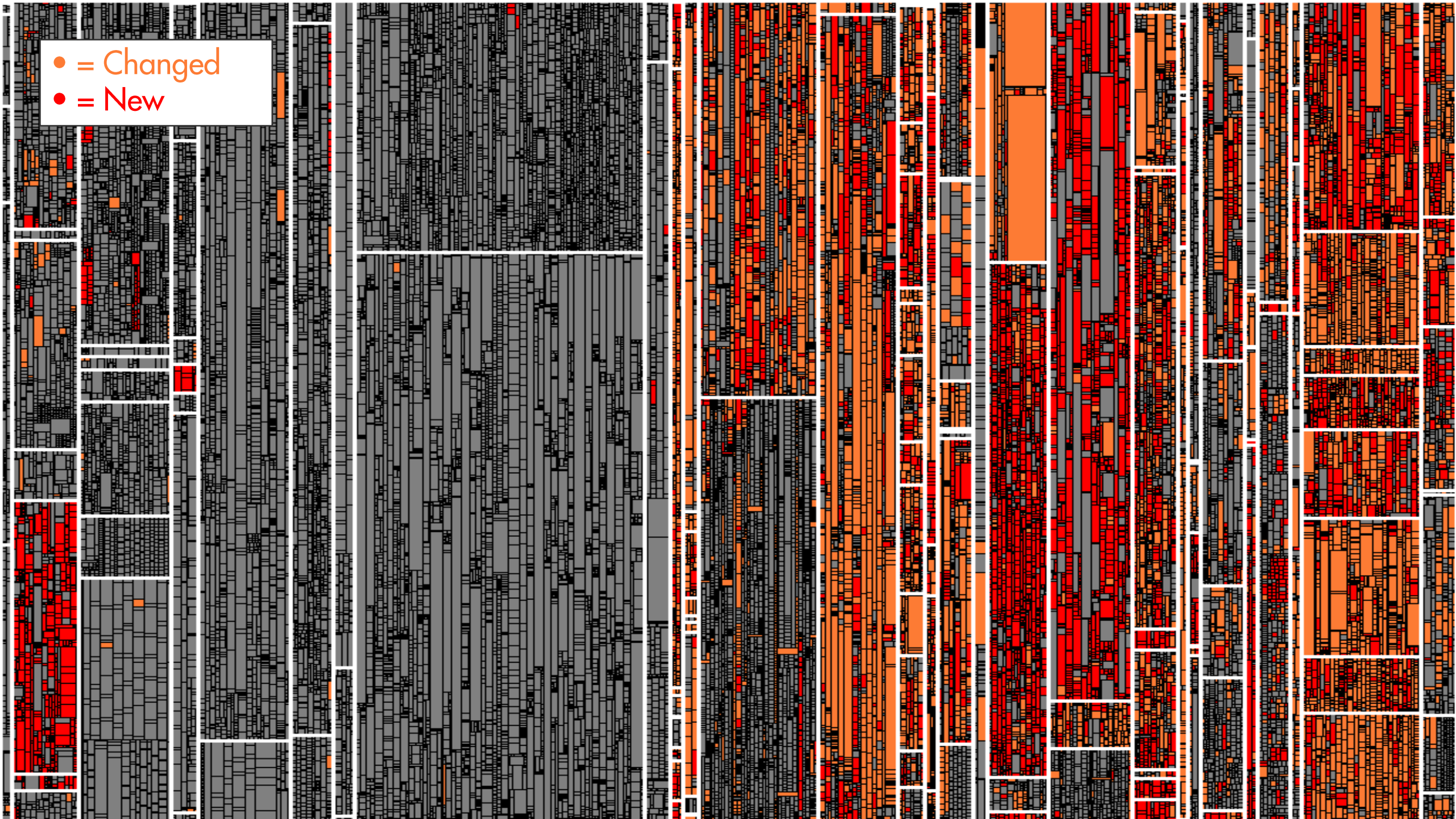
» Wo sind Fehler im Code? «

» Welche Tests sollte ich jetzt ausführen? «

» Wo sind ungetestete Änderungen «

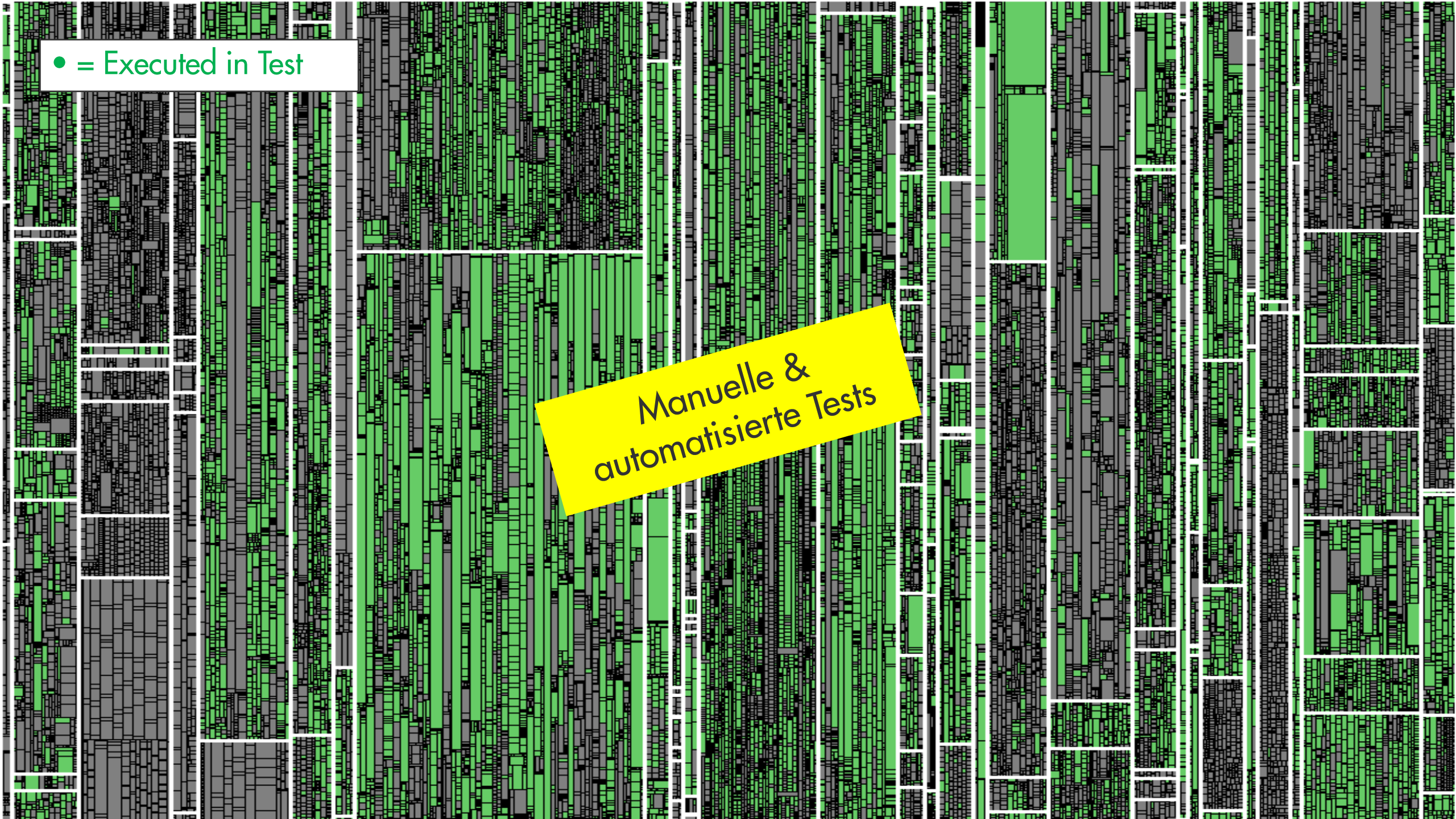


- = Changed
- = New

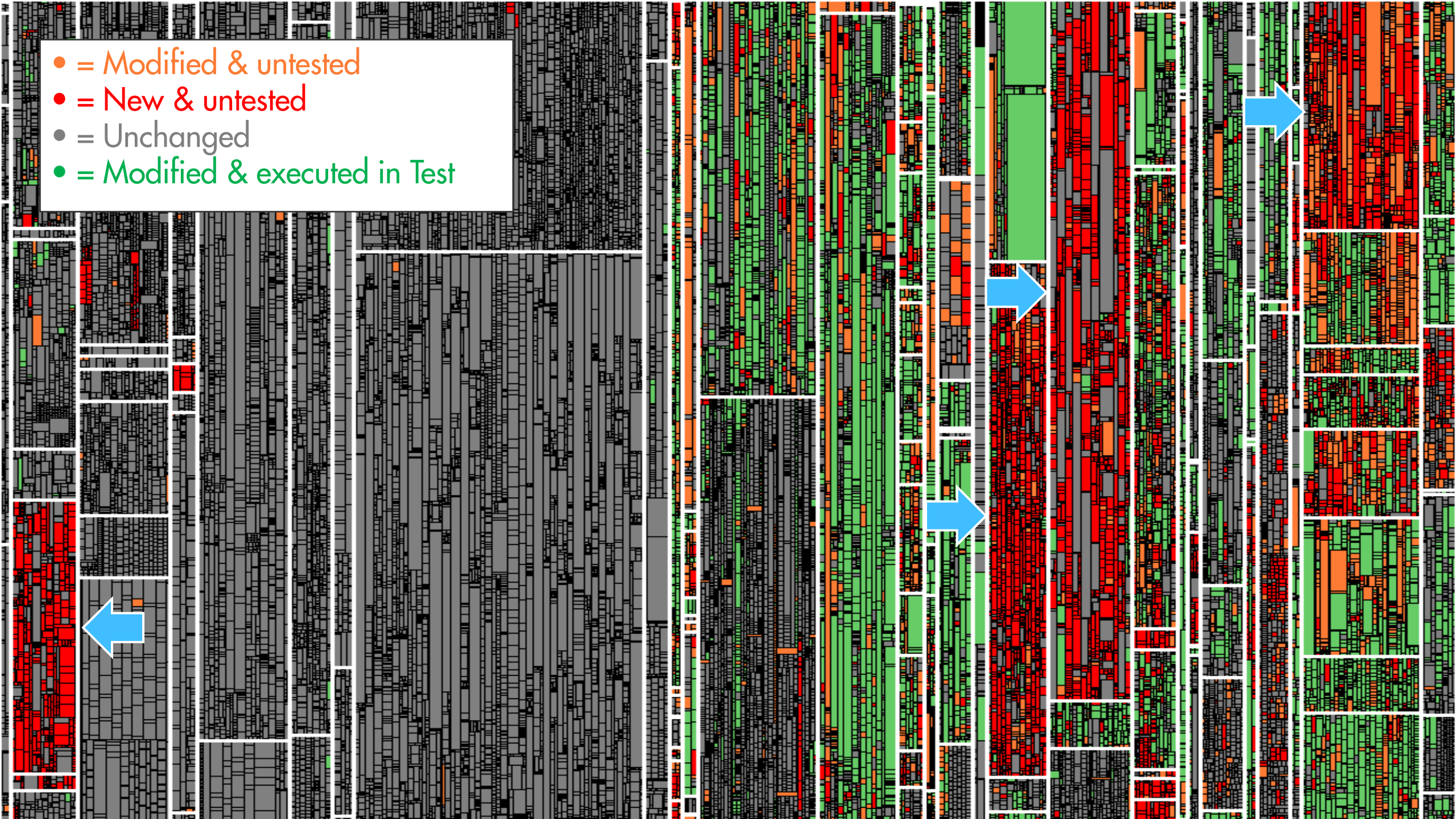



















● = Executed in Test





Manuelle &
automatisierte Tests




- = Modified & untested
- = New & untested
- = Unchanged
- = Modified & executed in Test




Issue # ▼	Subject	Done		Test Gap
🔗 TS-10549	Undo/Redo for web-based architecture editor	Done		0% 
🔗 TS-10784	Fix long method finding in TaintAnalysisRunner	Done		0% 
🔗 TS-10923	Implement metric 'Nesting Depth' for Simulink	Done		29% 
🔗 TS-11364	External findings are not registered during first upload	Done		14% 
🔗 TS-11942	Manual test coverage upload during development	Done		43% 
🔗 TS-12050	Tool for transferring findings blacklists and tasks	Done		50% 
🔗 TS-12262	Cannot set or alter alias without reanalysis	Done		0% 
🔗 TS-13151	Fetch parent relationship of TFS work items	Done		0% 

Issue # ▾	Subject		Test Gap
TS-14421	Get rid of TestGapSynchronizer block	Done 	0% 
TS-14733	Remove Dataflow blocks	Done 	22% 

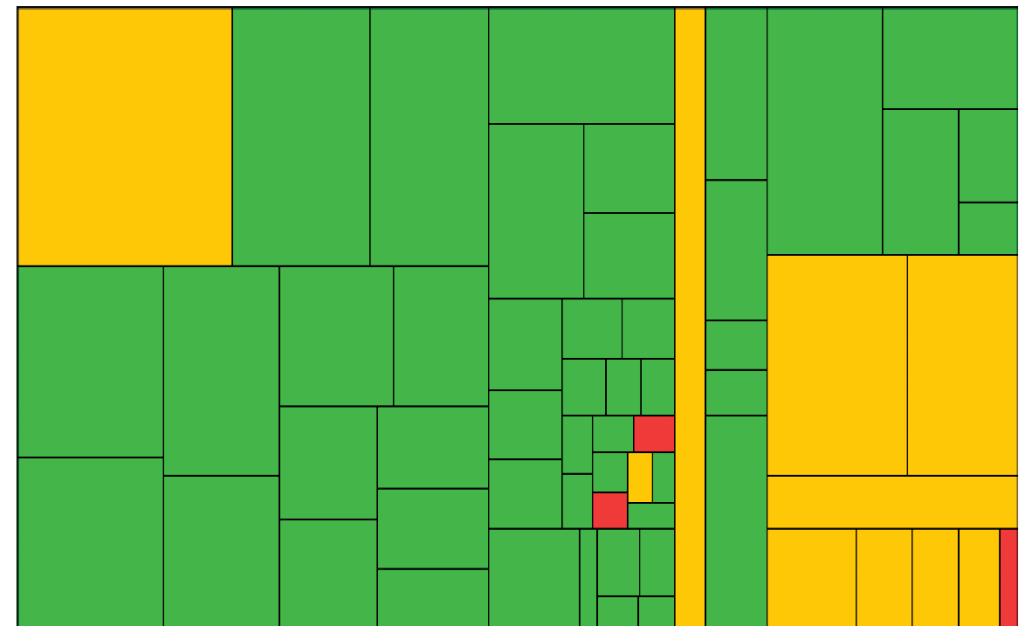
Done Issue TS-14733 - Remove Dataflow blocks

Creator:  (on Apr 06 2018 19:44) Last update: Aug 24 2018 09:32

Assignee: 

Project	Type	Priority	Resolution	Fix Version
TS	Maintenance	Normal	Green	Teamscale 4.5
Component	Labels	Affected Version	Customer	Customer Issue
Backend	Performance			
Epic Name	Freshdesk URL	Merge Request		
		https://git.cqse.eu/cqse/teamscale/3621		

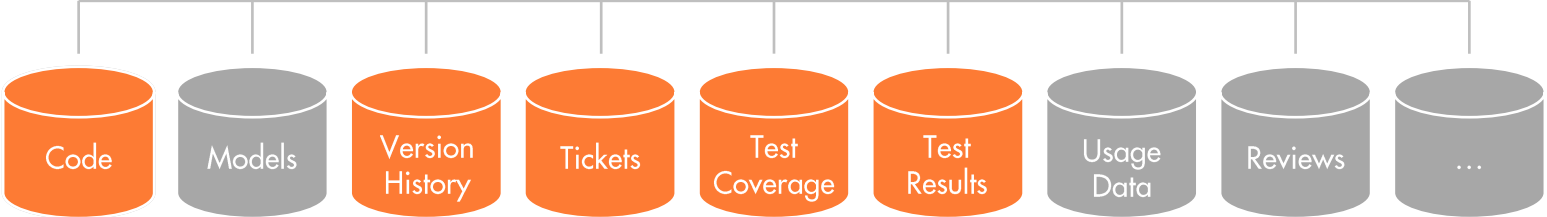
Aug 15 2018 12:37–Now | Test Gap: 22%

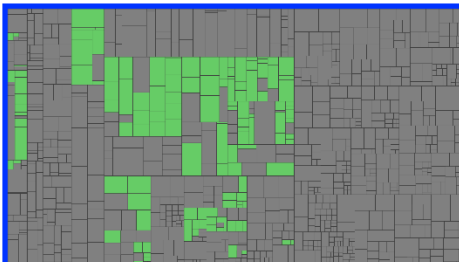
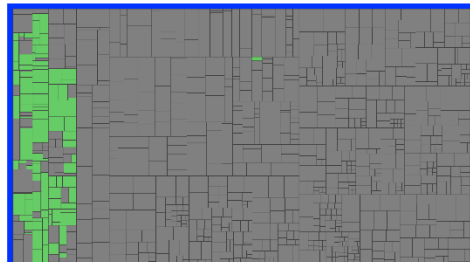
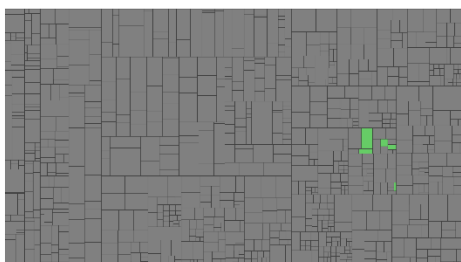
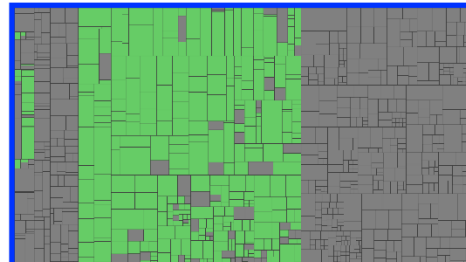
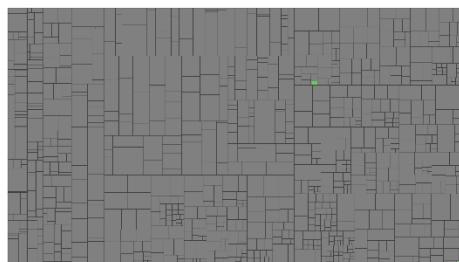
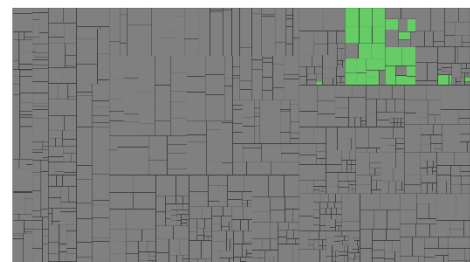
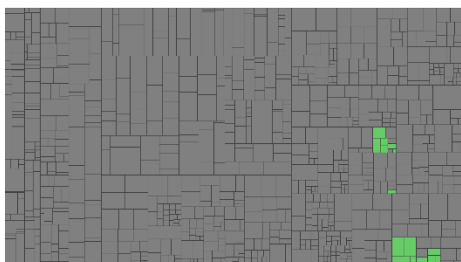
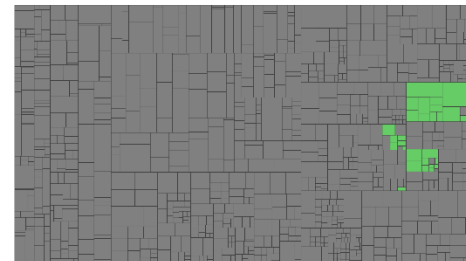
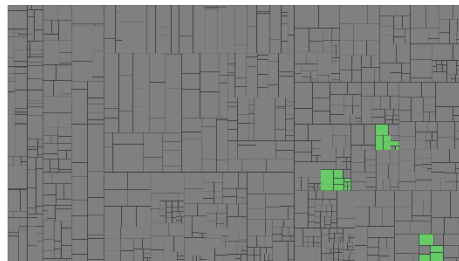
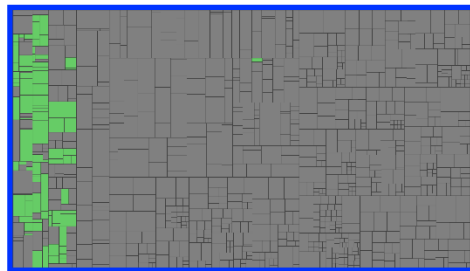
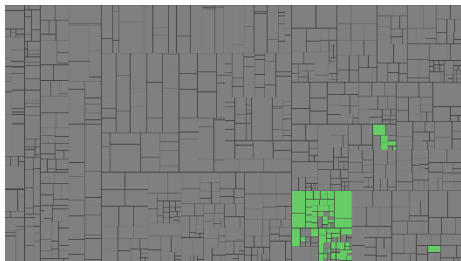
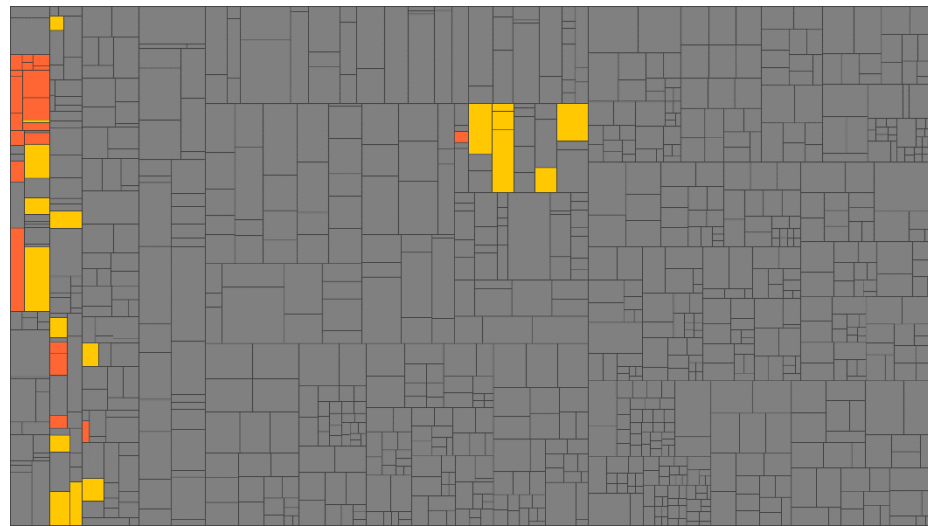


» Wo sind Fehler im Code? «

» Welche Tests sollte ich jetzt ausführen? «

» Wo sind ungetestete Änderungen «





... 4000+ test cases
don't execute
changed code

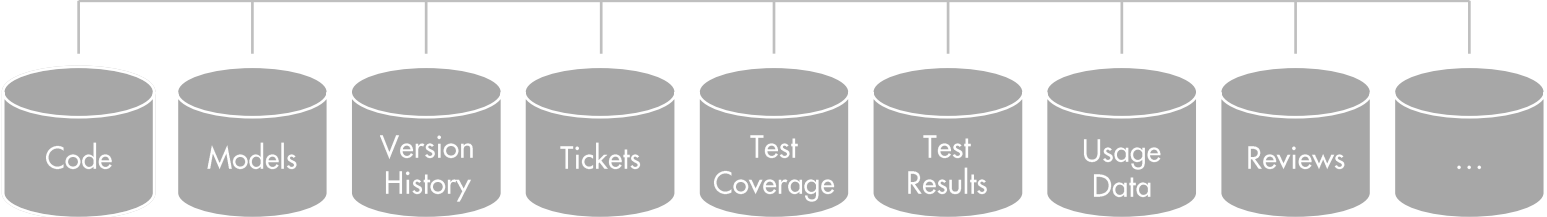
» Wo sind Fehler im Code? «

» Welche Tests sollte ich jetzt ausführen? «

» Wo sind ungetestete Änderungen «



Software Intelligence

Kontakt – Ich freue mich auf Fragen 😊



Dr. Elmar Jürgens

juergens@cqse.eu

+49 179 675 3863

CQSE GmbH
Centa-Hafenbrädl-Str 59
81249 München
www.cqse.eu

