Test Intelligence

für Architektinnen und Architekten

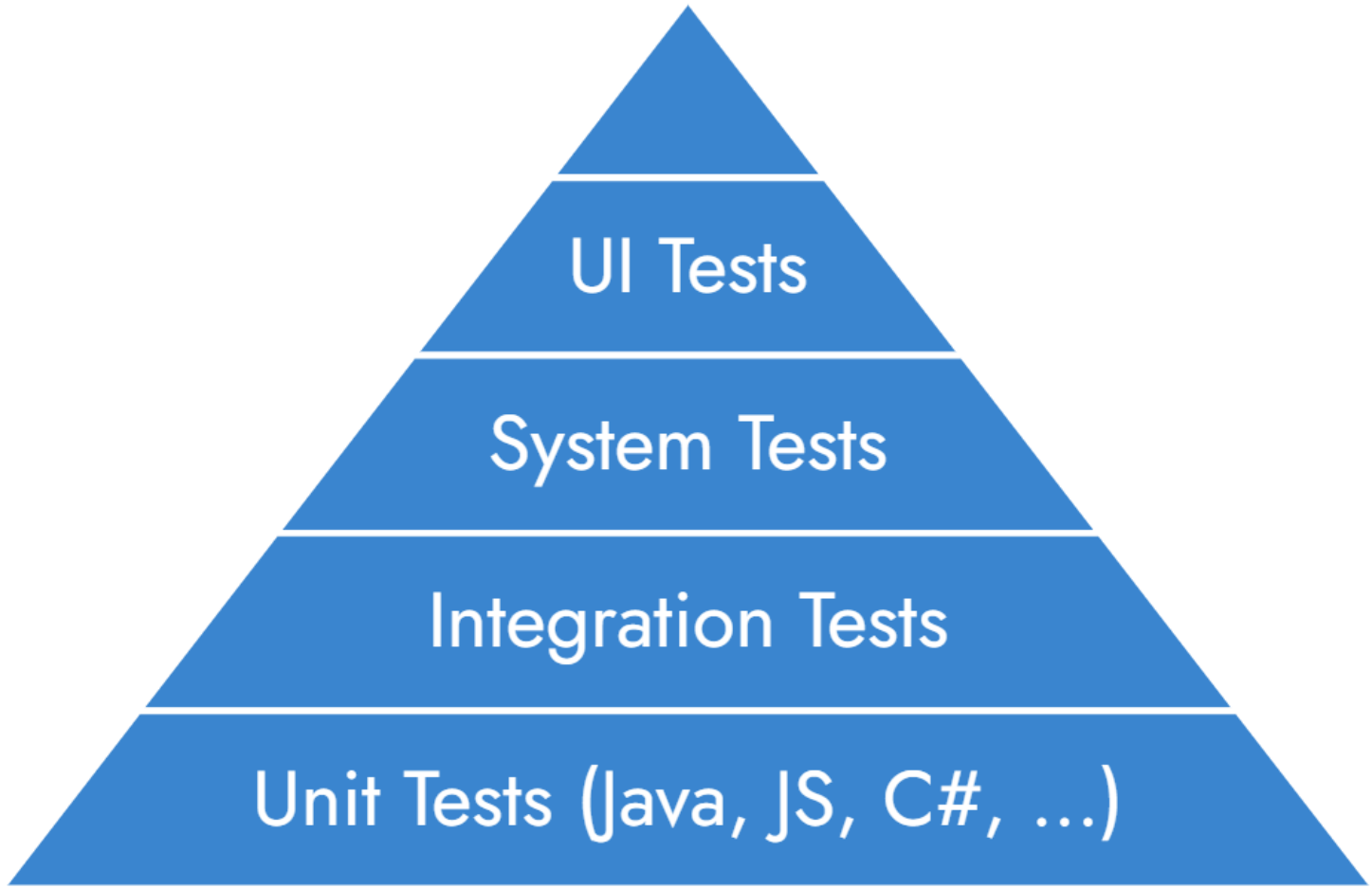
Dr. Elmar Jürgens

CQSE

Slower
More Flaky

Total Count
13687

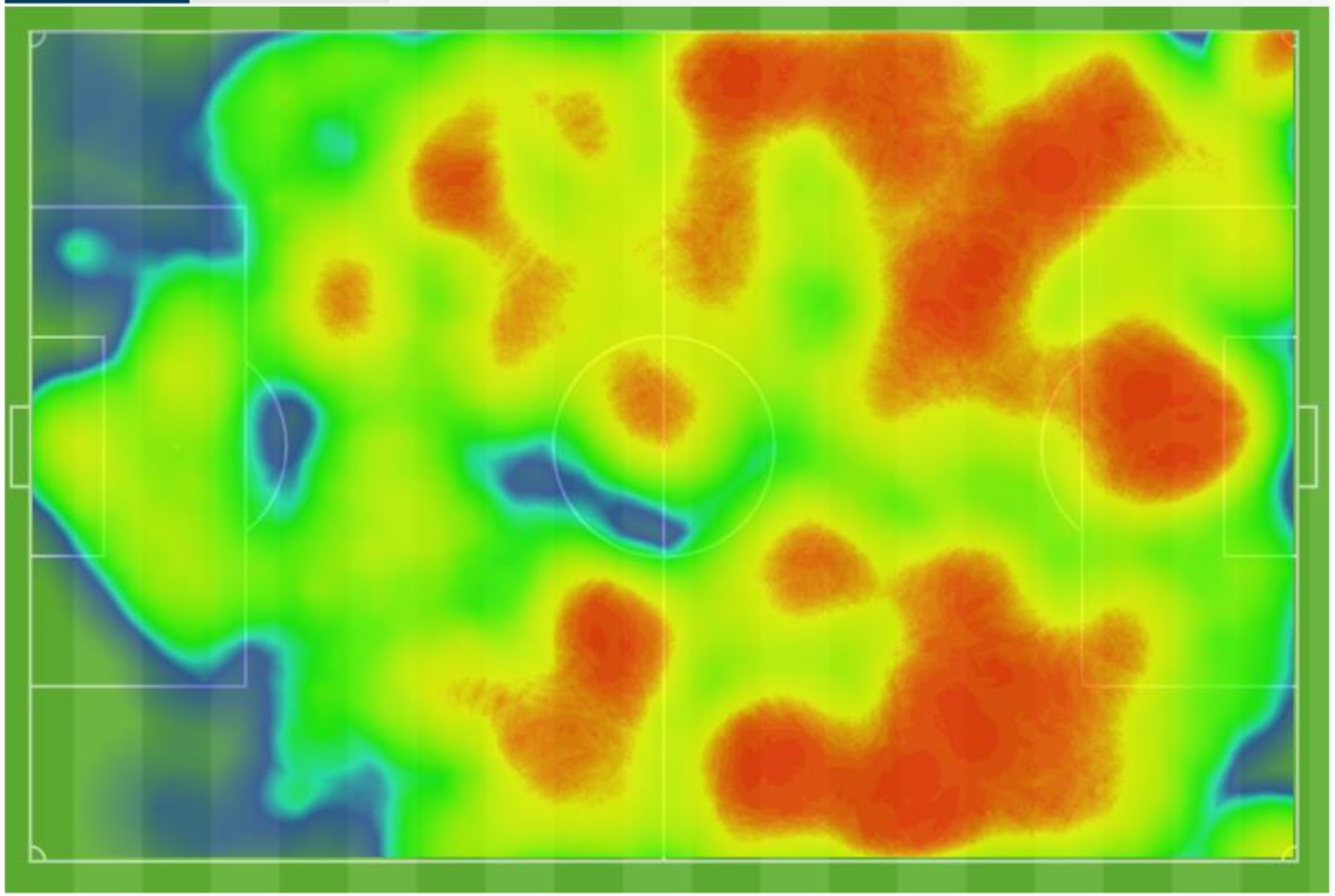
Total Duration
~11,4h



Faster
More stable

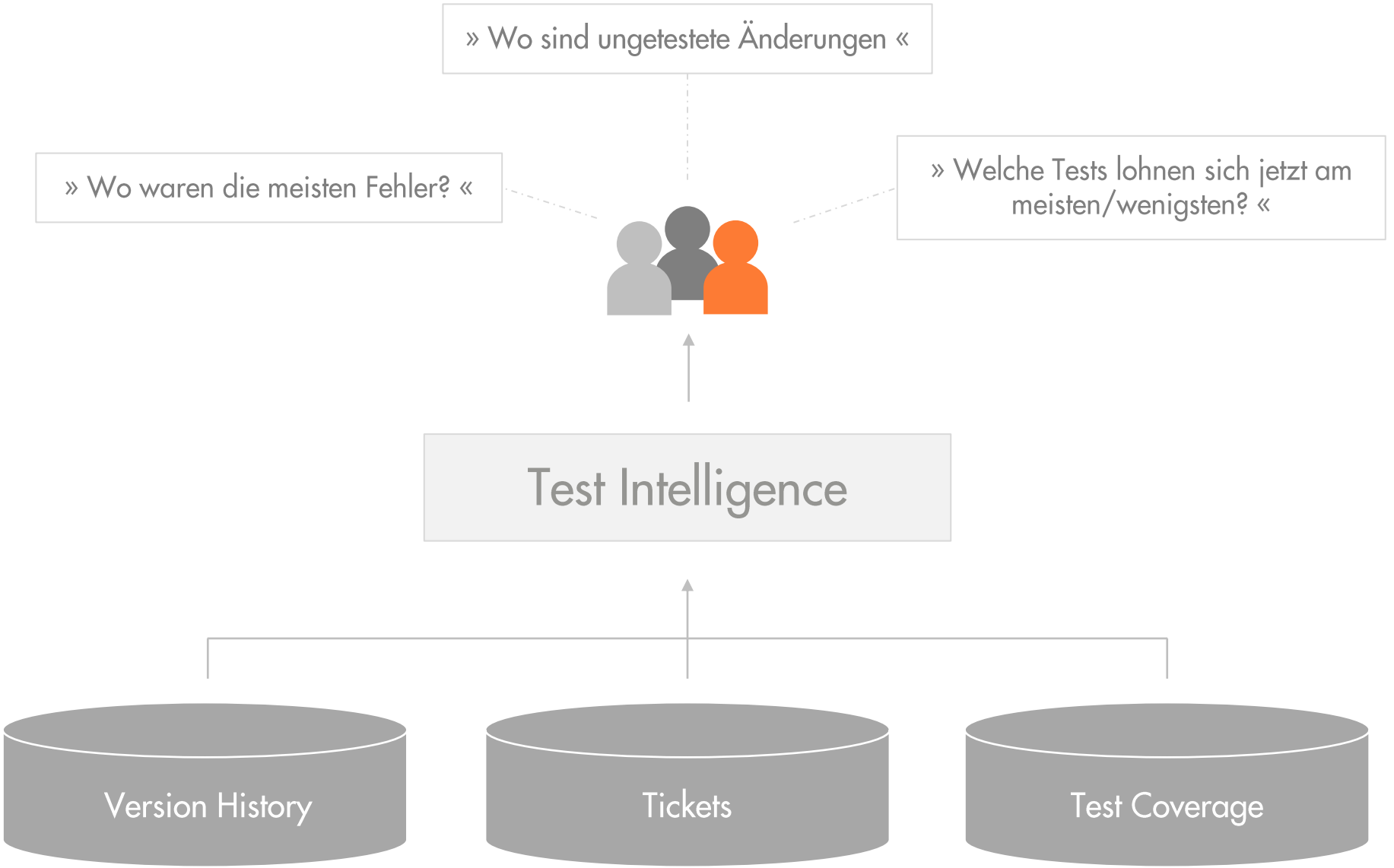
State: April 2023

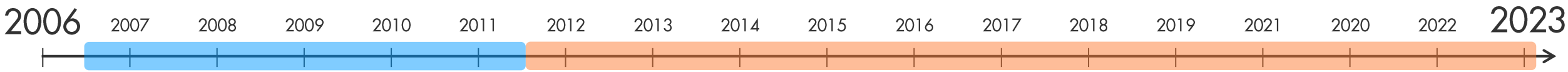
Heatmap | Touchmap

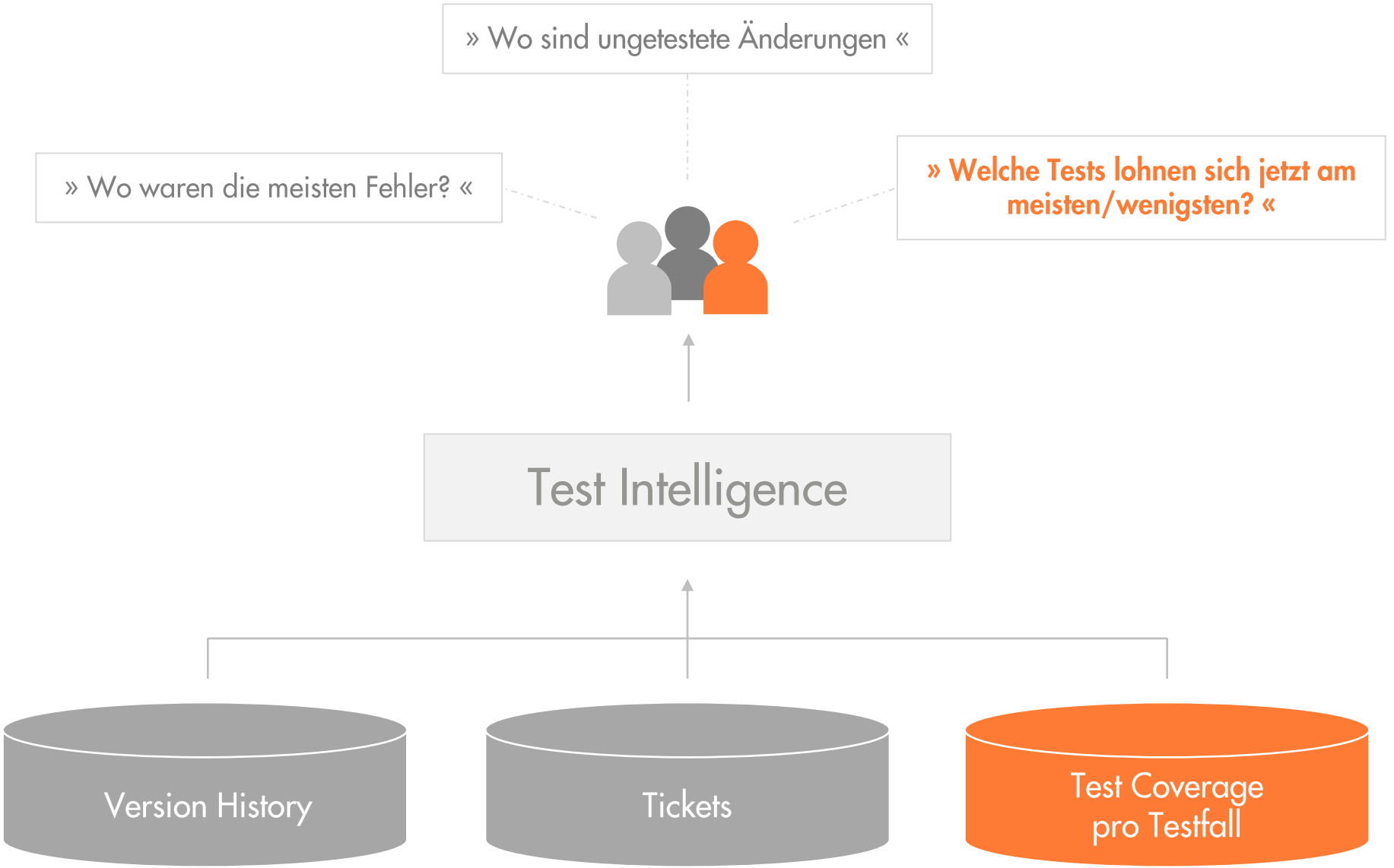


FC Bayern München			Paris Saint-Germain		
FC Bayern München			Paris Saint-Germain		
TW	1	Manuel Neuer	TW	1	Keylor Navas
AW	21	Lucas Hernández	AW	5	Marquinhos
AW	5	Benjamin Pavard	AW	31	Colin Dagba
AW	4	Niklas Süle	AW	22	Abdou Diallo
AW	27	David Alaba	AW	3	Presnel Kimpembe
MF	29	Kingsley Coman	MF	23	Julian Draxler
MF	10	Leroy Sané	MF	27	Idrissa Gueye
MF	25	Thomas Müller	MF	10	Neymar
MF	6	Joshua Kimmich	MF	15	Danilo Pereira
MF	18	Leon Goretzka	MF	11	Ángel Di María
ST	13	Eric Maxim Choupo-Moting	ST	7	Kylian Mbappé
Einwechselspieler			Einwechselspieler		
17		Jérôme Boateng	25		Mitchel Bakker
19		Alphonso Davies	18		Moise Kean
			21		Ander Herrera
			12		Rafinha

https://sportdaten.welt.de/fussball/champions-league/bayern-muenchen/paris-st-germain/2021-04-07/clvfh/2021/142784/heatmap


» Wo sind ungetestete Änderungen «
» Wo waren die meisten Fehler? «
» Welche Tests lohnen sich jetzt am meisten/wenigsten? «
Test Intelligence
Version History
Tickets
Test Coverage


2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2021
2020
2022
2023


» Wo sind ungetestete Änderungen «
» Wo waren die meisten Fehler? «
» Welche Tests lohnen sich jetzt am meisten/wenigsten? «
Test Intelligence
Version History
Tickets
Test Coverage pro Testfall

Picture1.png - Pixelitor 4.2.3

File Edit Layer Select Image Color Filter View Help

Sample Only the Active Layer/Mask

Untitled1 × Picture1.png ×

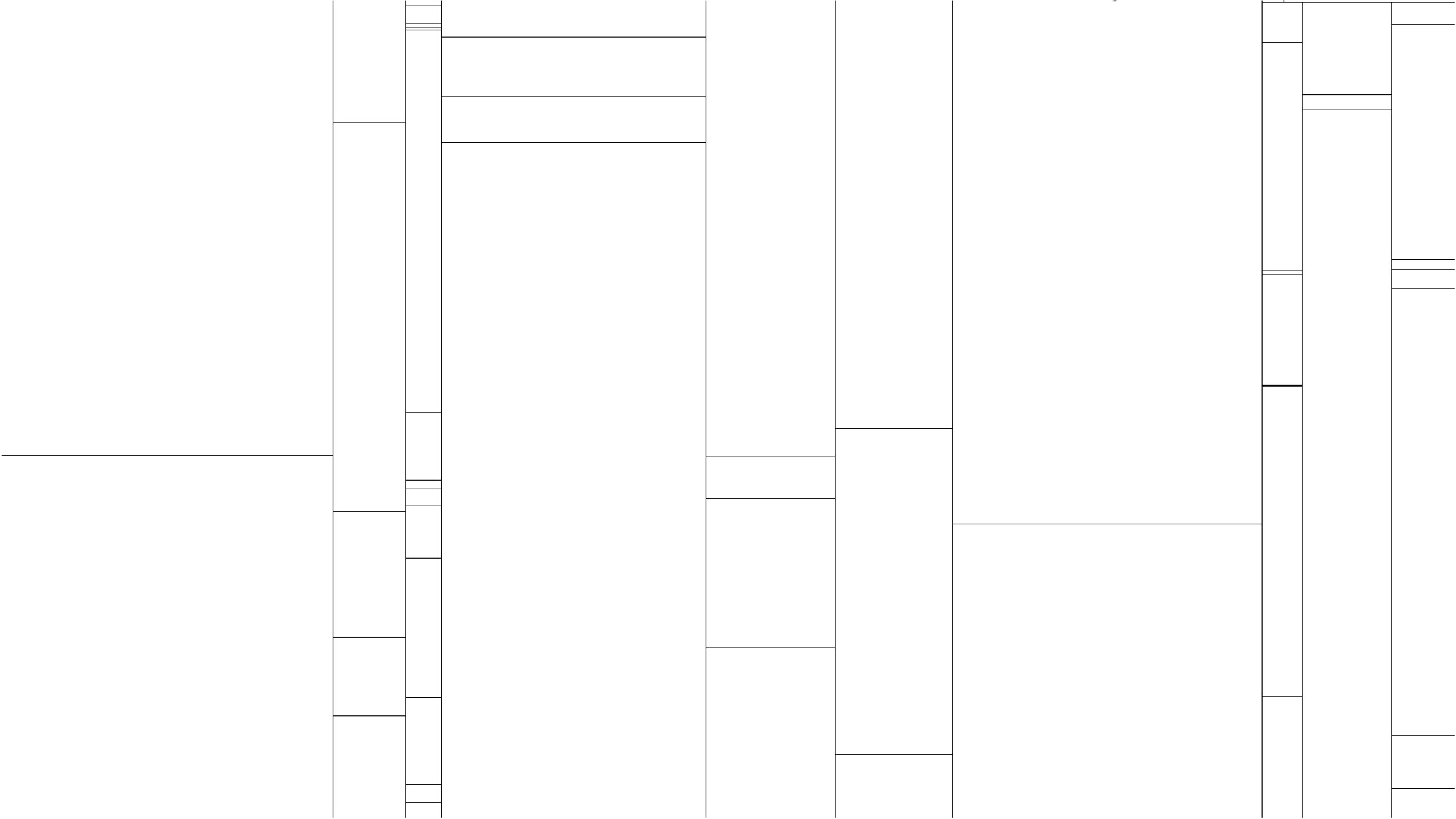
Layers

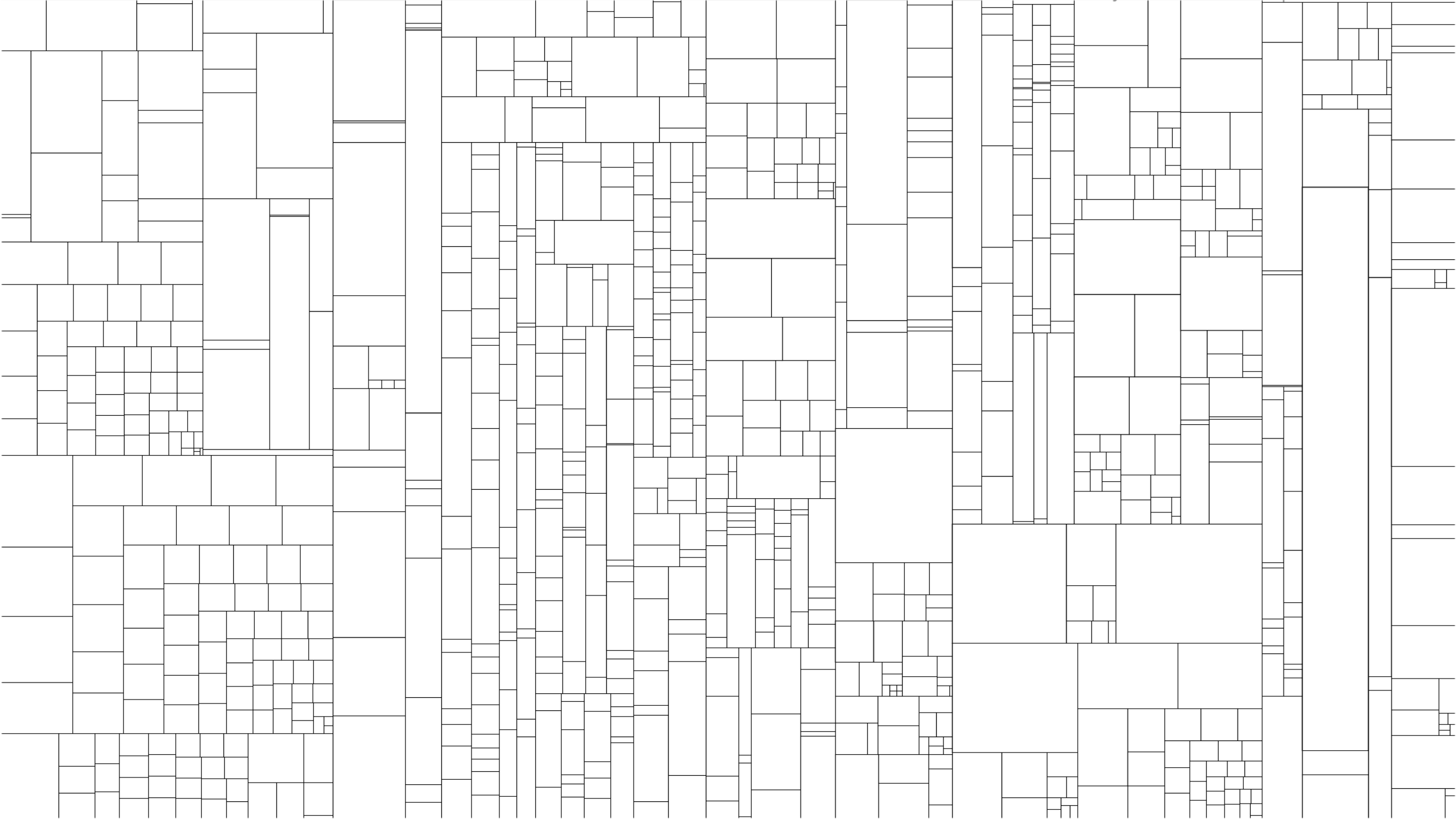
Opacity: 100 % Normal

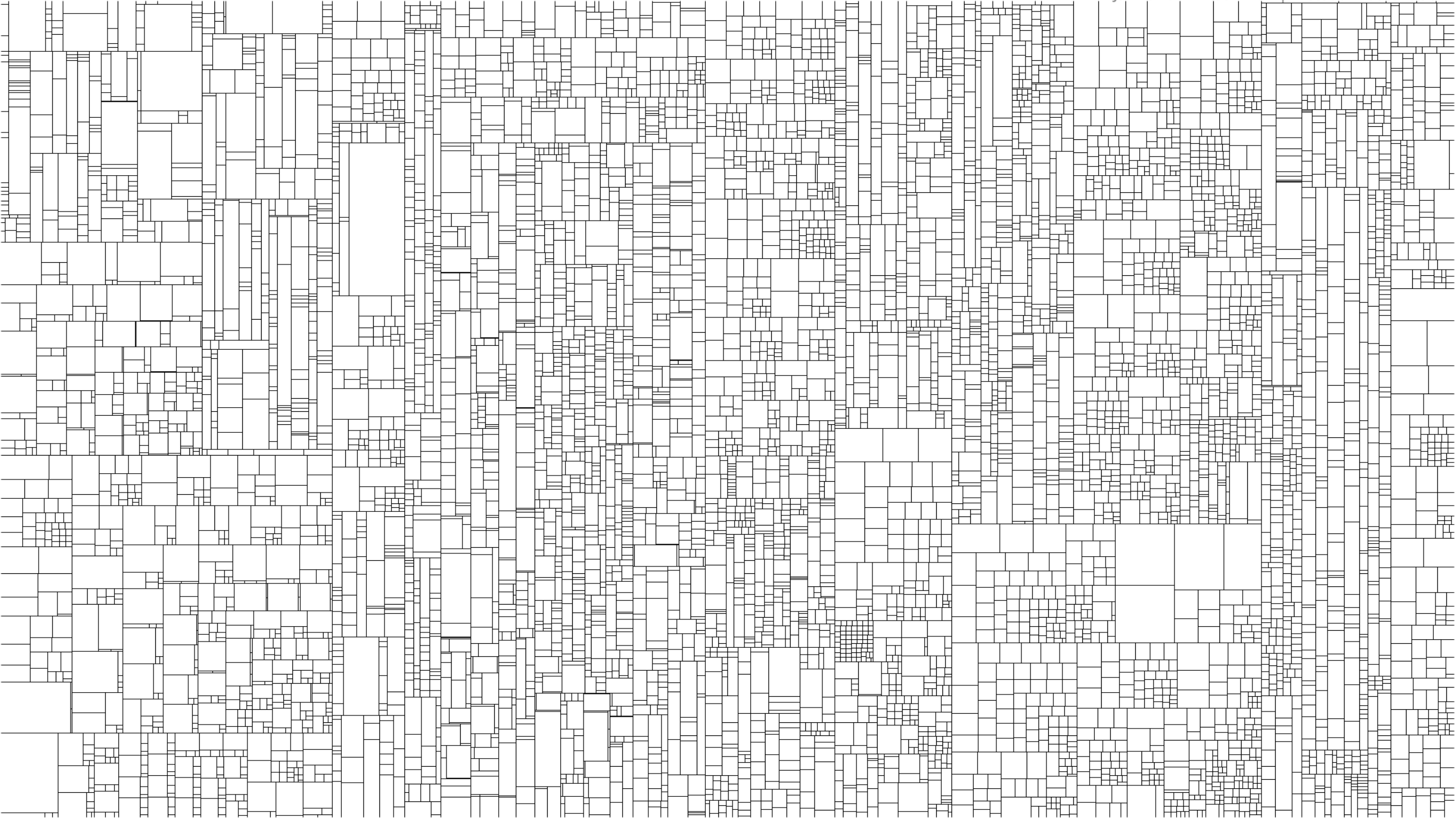
layer 1

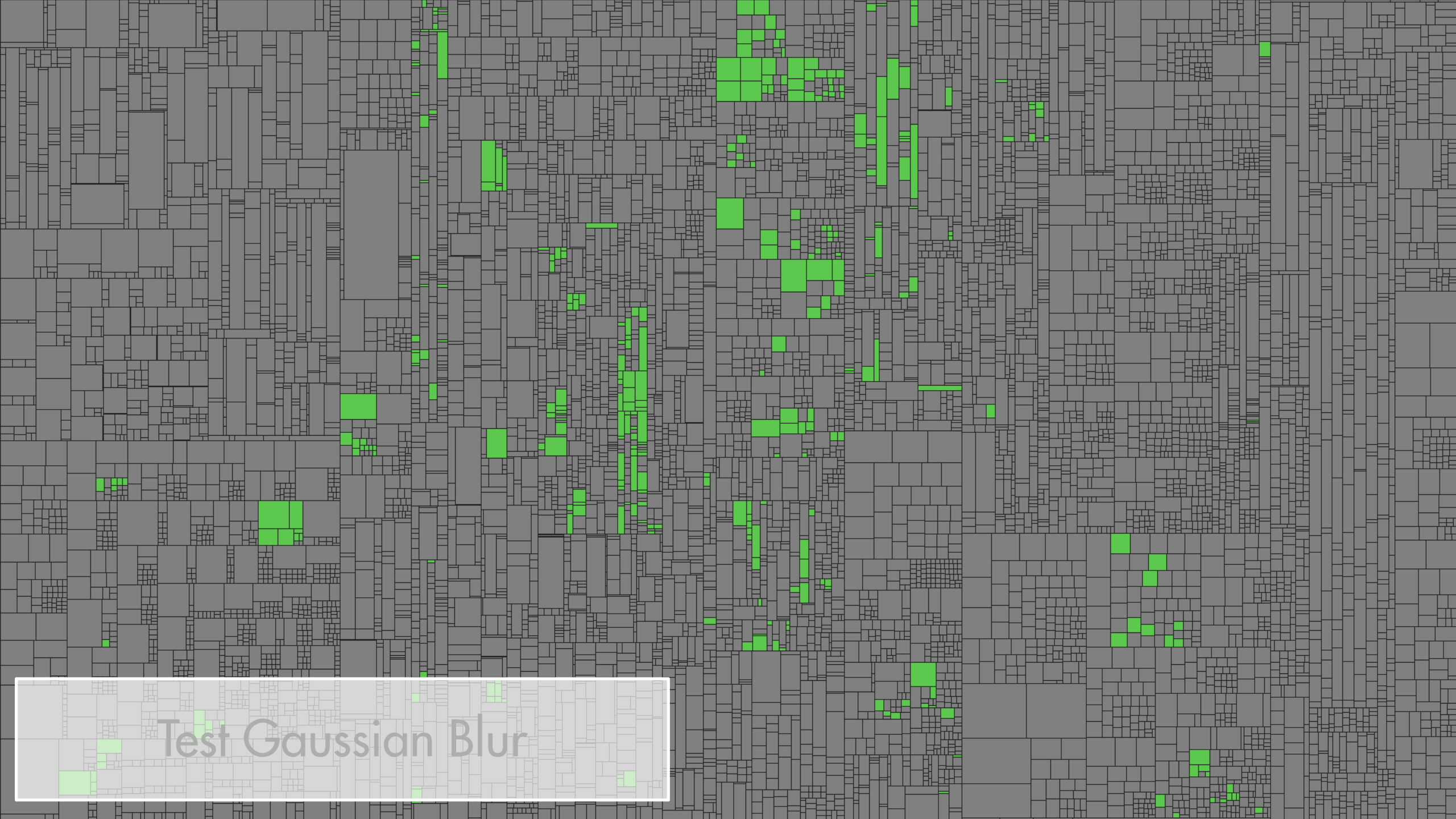
Picture1.png was opened.

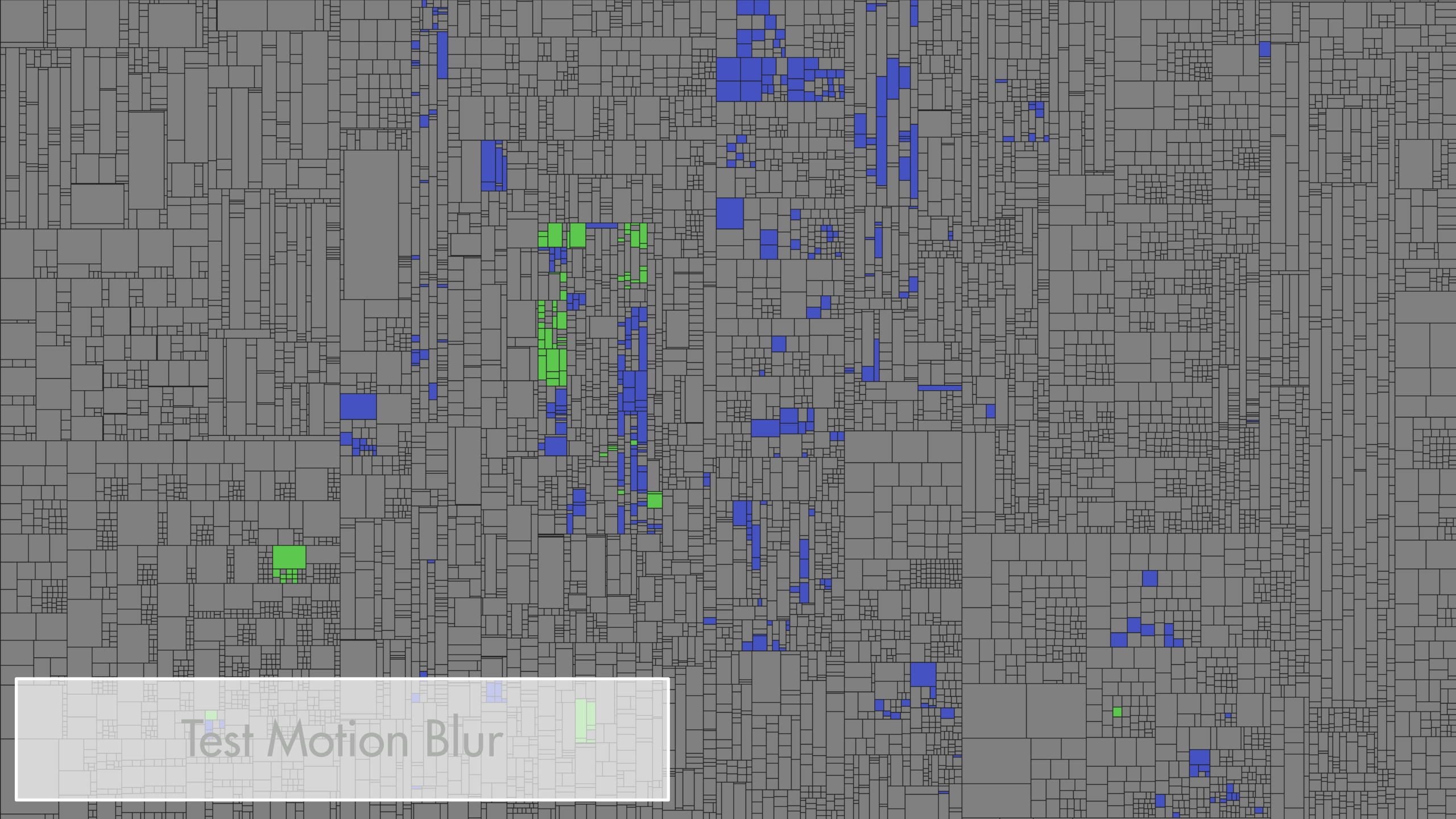
Zoom: 100% Fit 100%

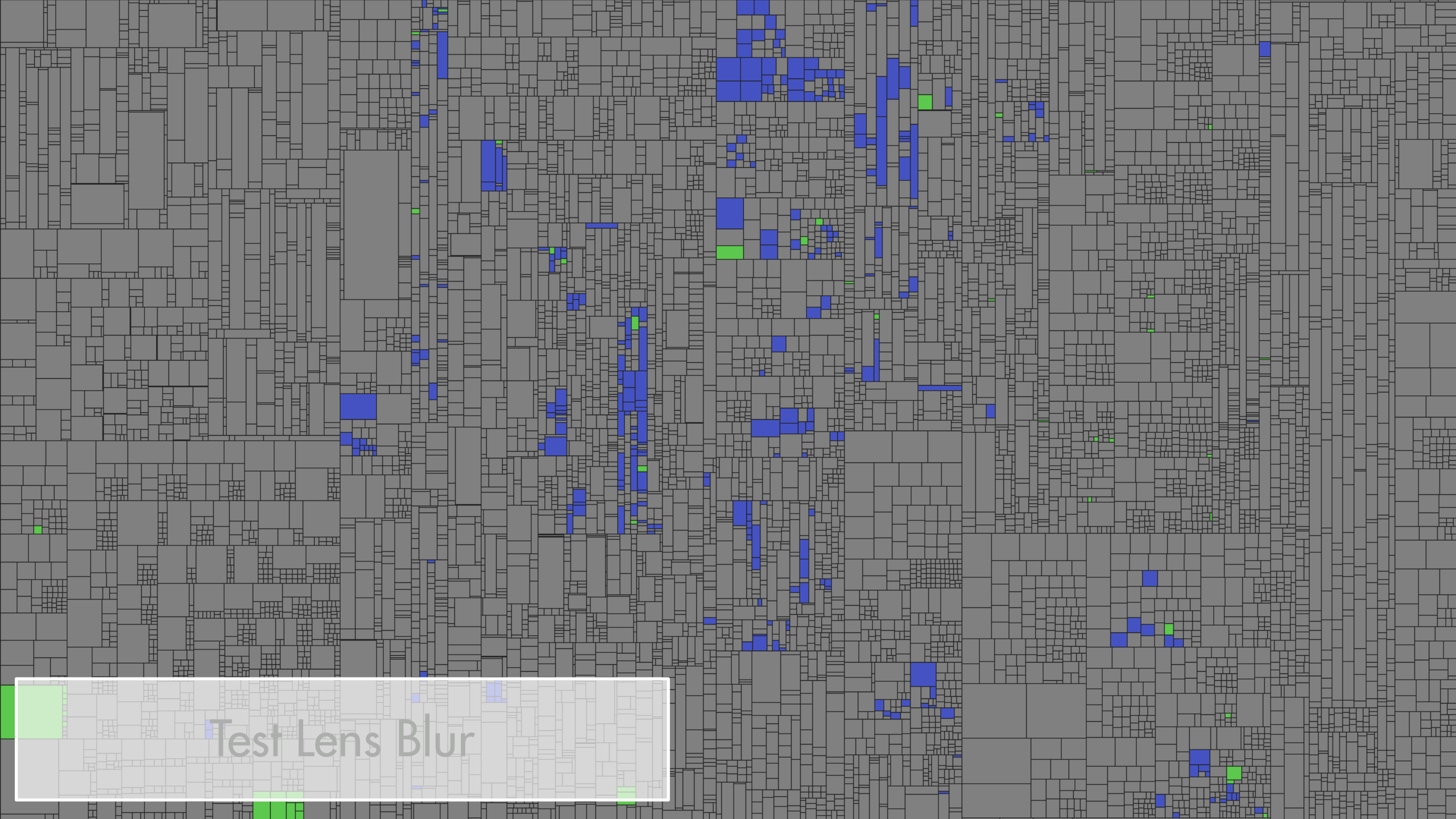


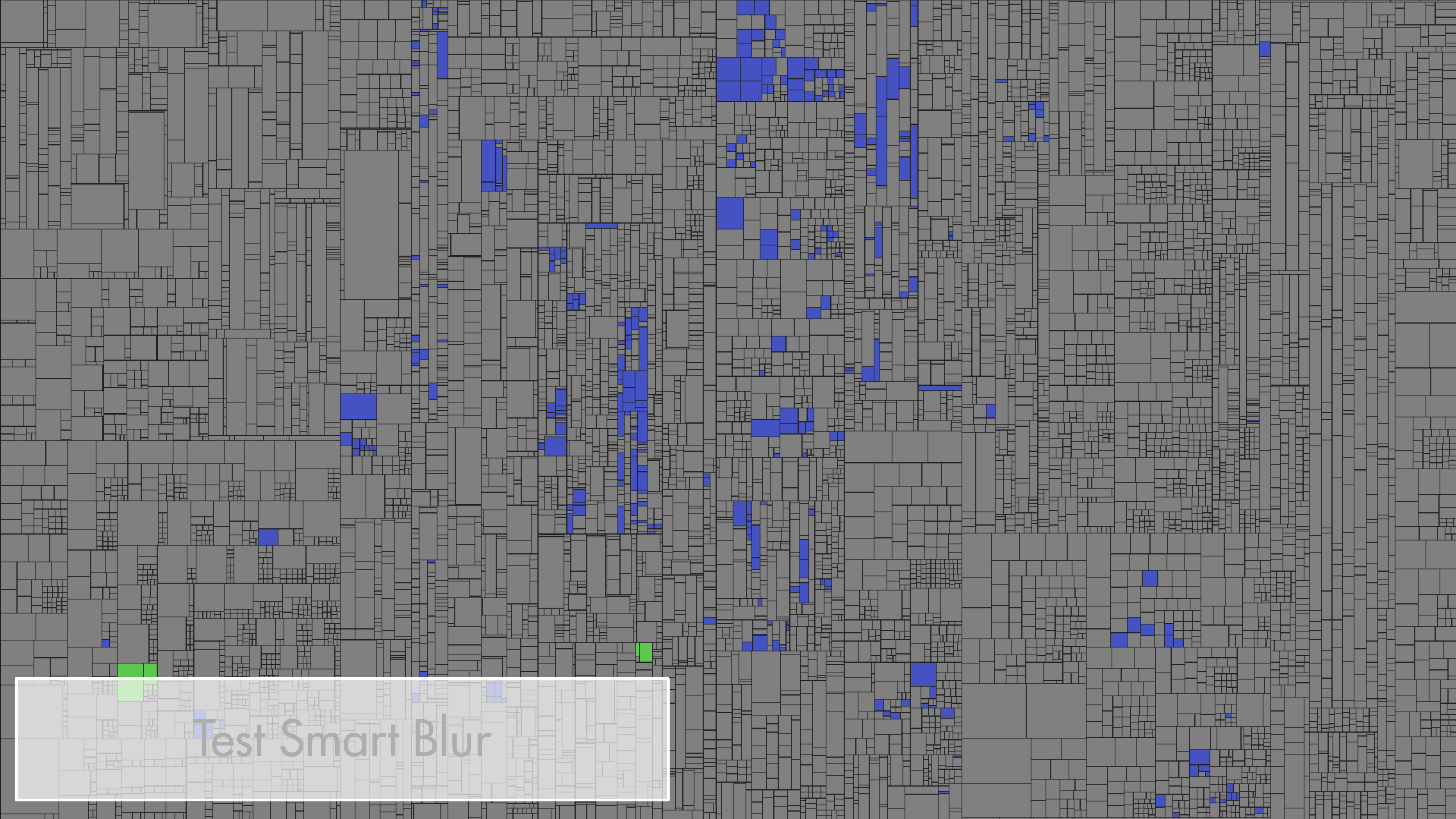


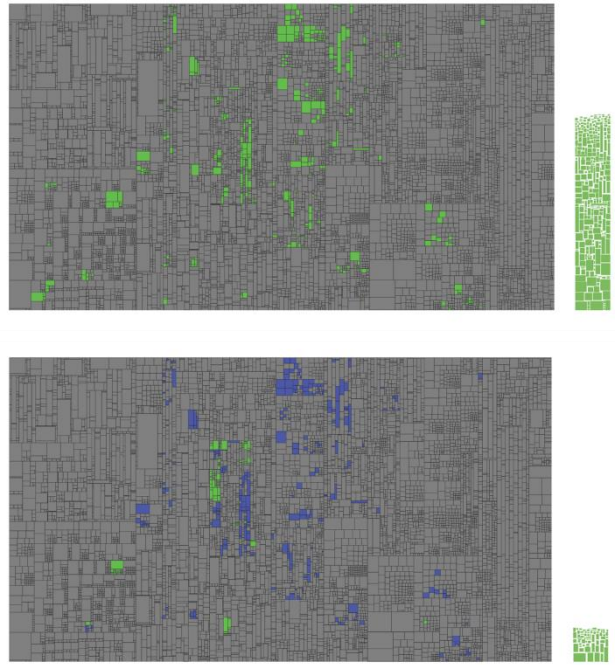



Test Gaussian Blur

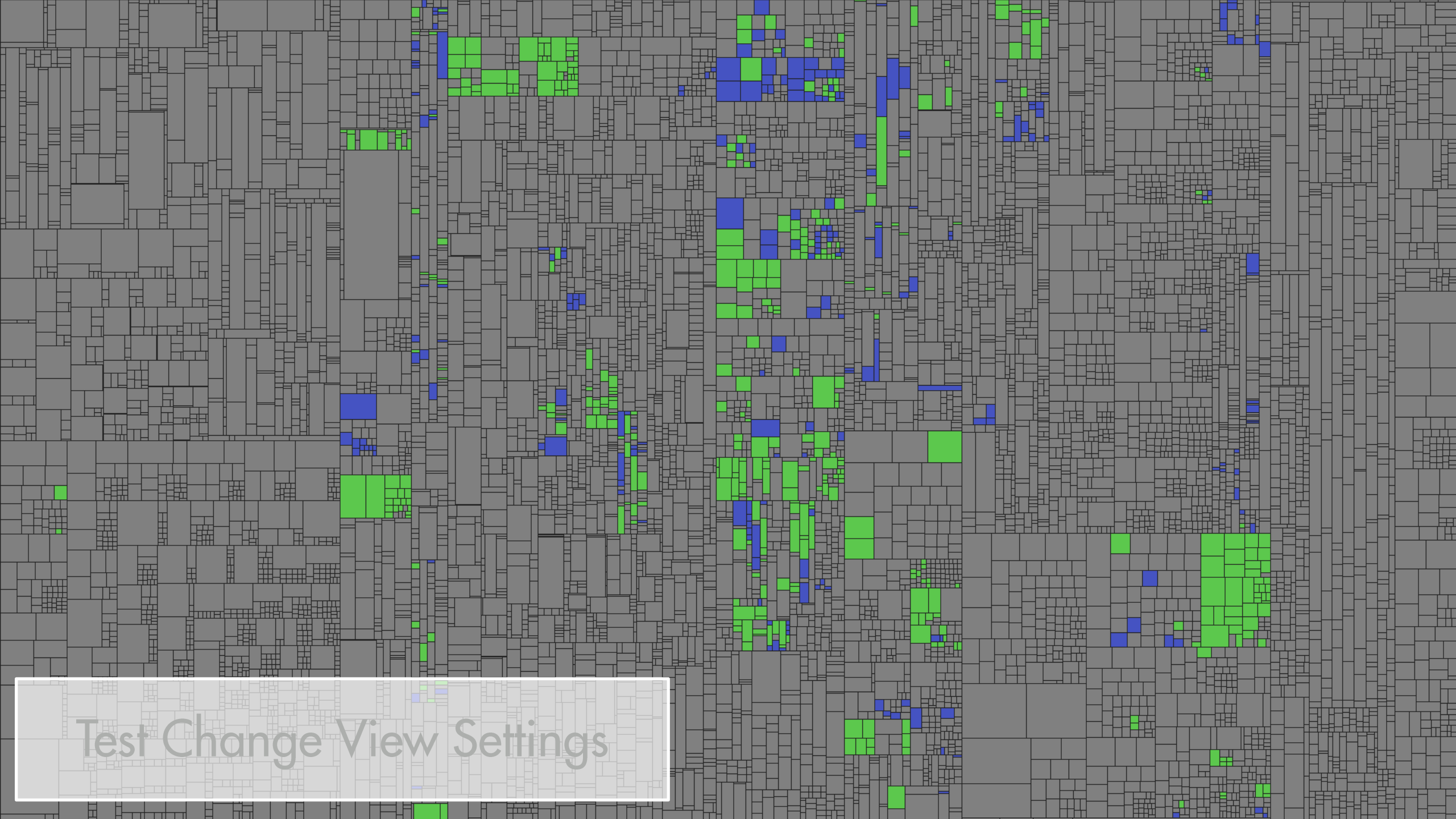

Test Motion Blur

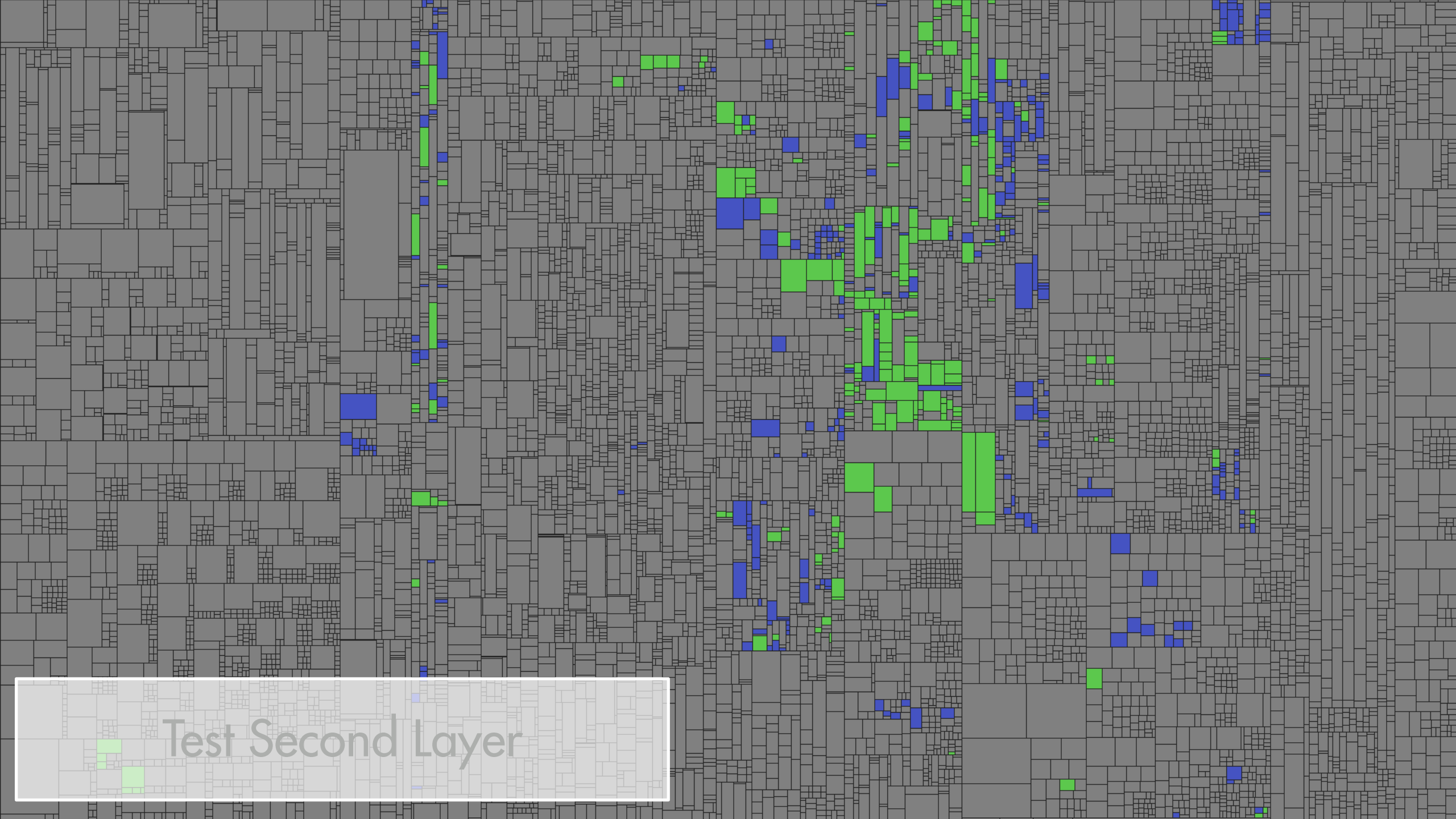

Test Lens Blur

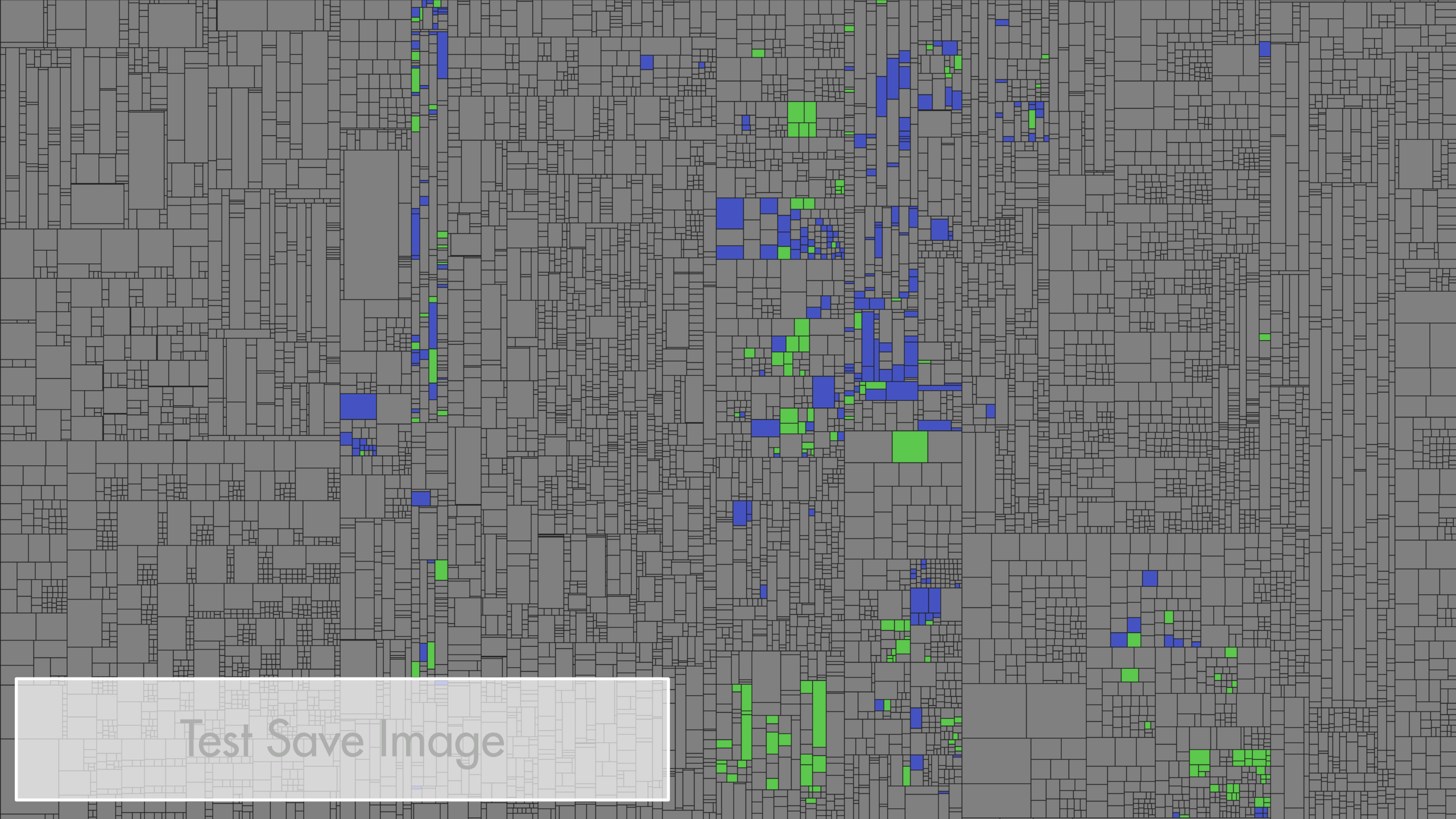

Test Smart Blur

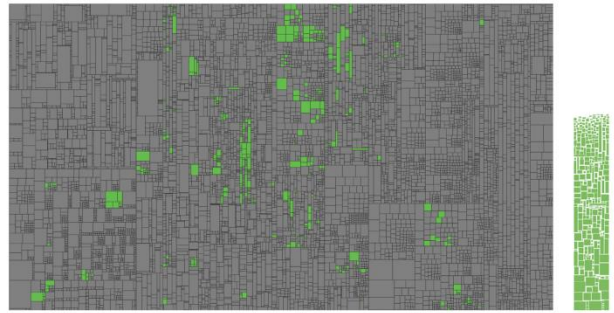


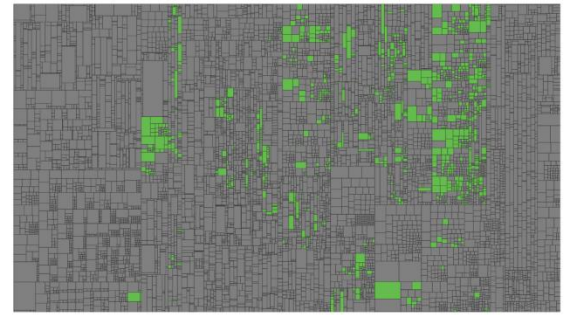

Test Create and Modify
Selection

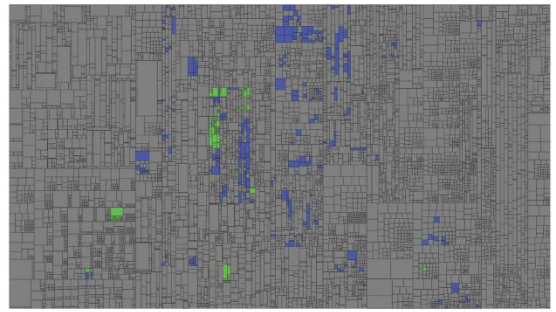

Test Change View Settings

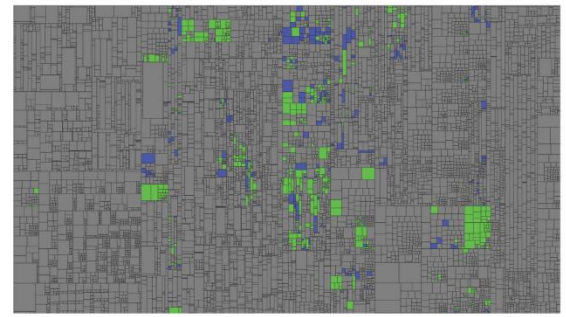

Test Second Layer

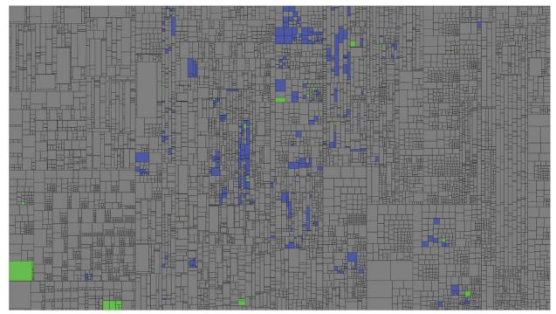

Test Save Image

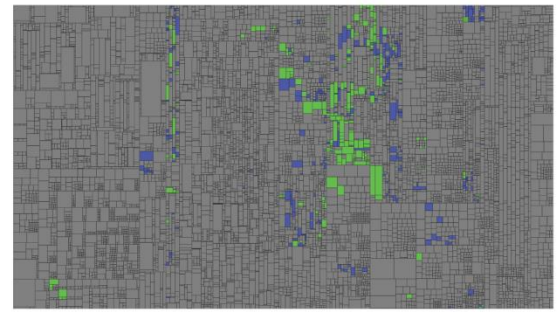


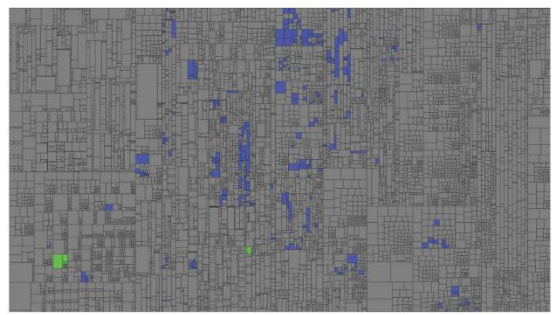


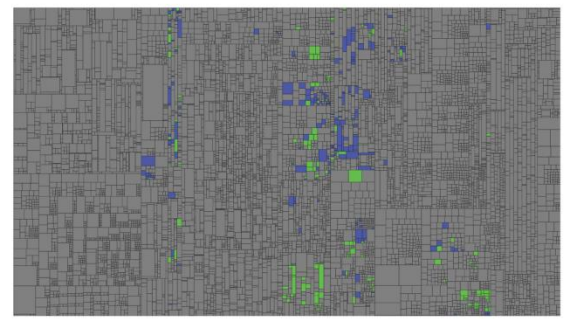



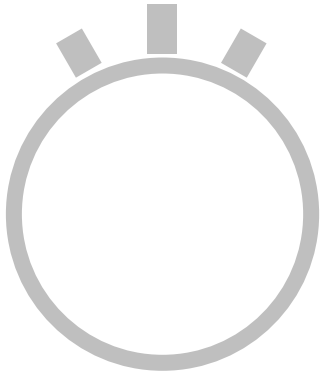


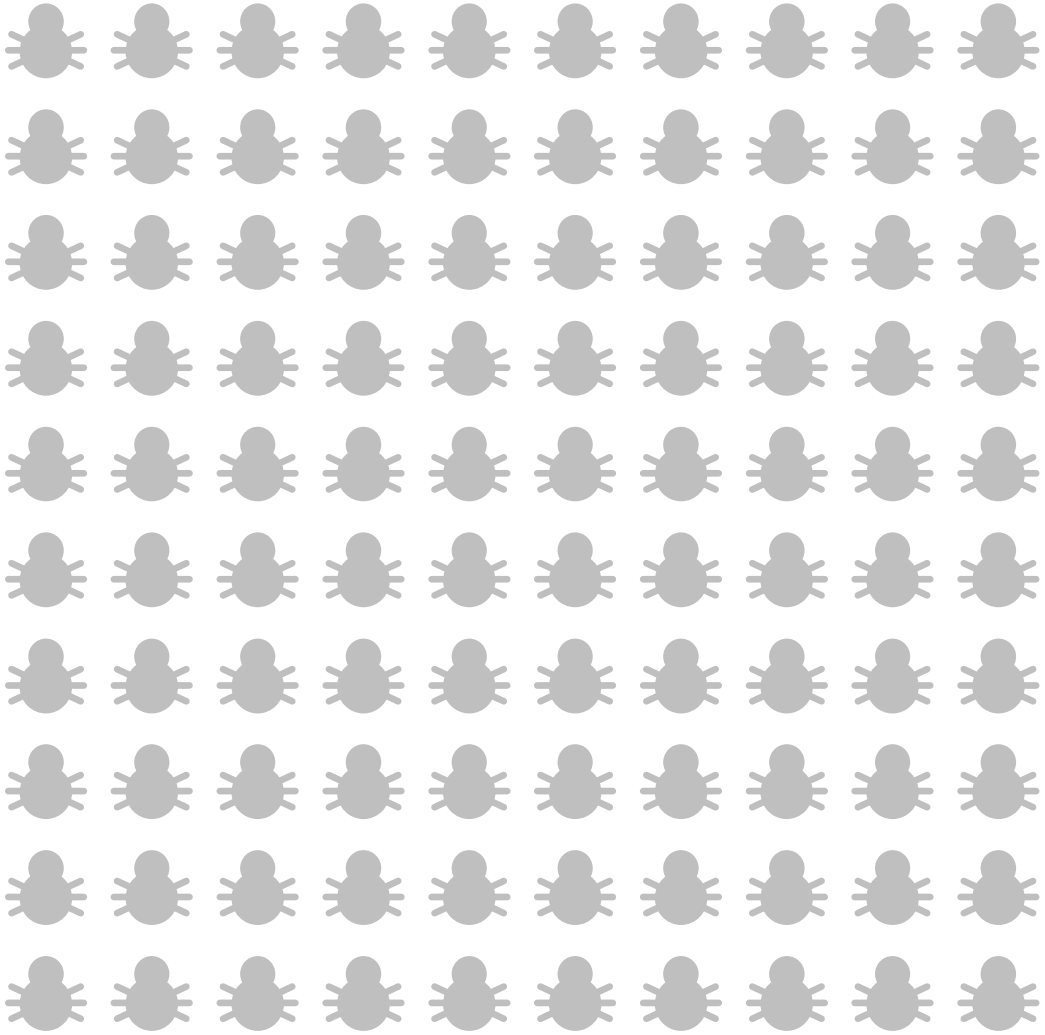




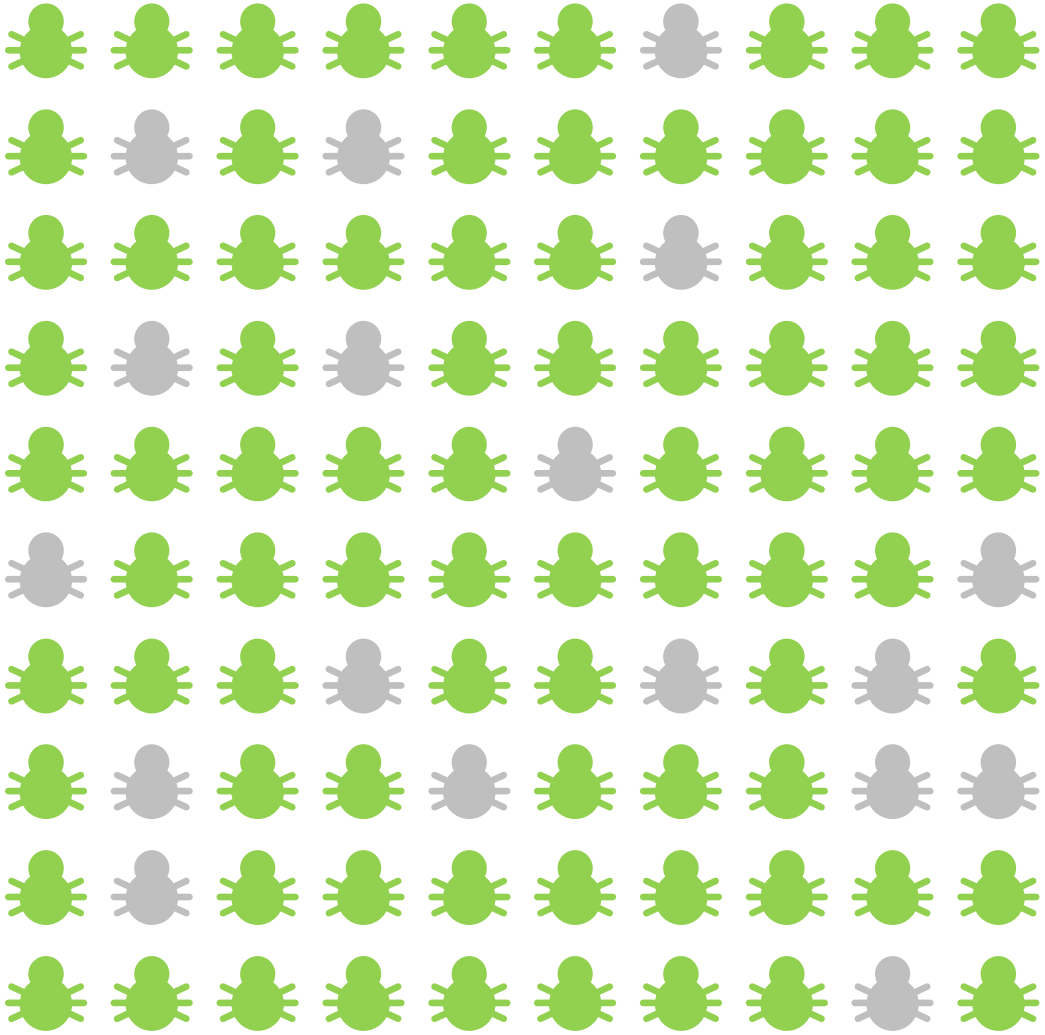




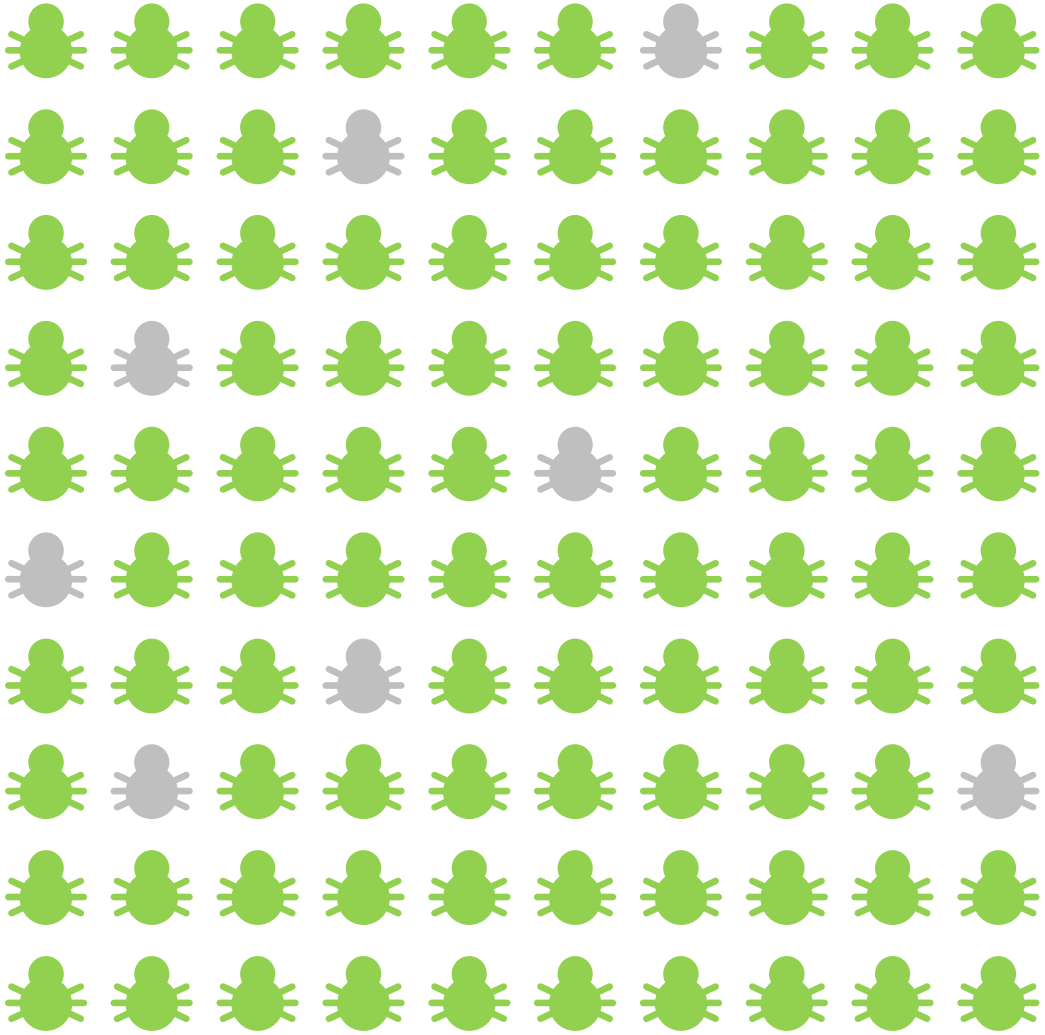





80 %



90 %



Pareto-Testliste in der CI

BVK Bayerische Versorgungskammer

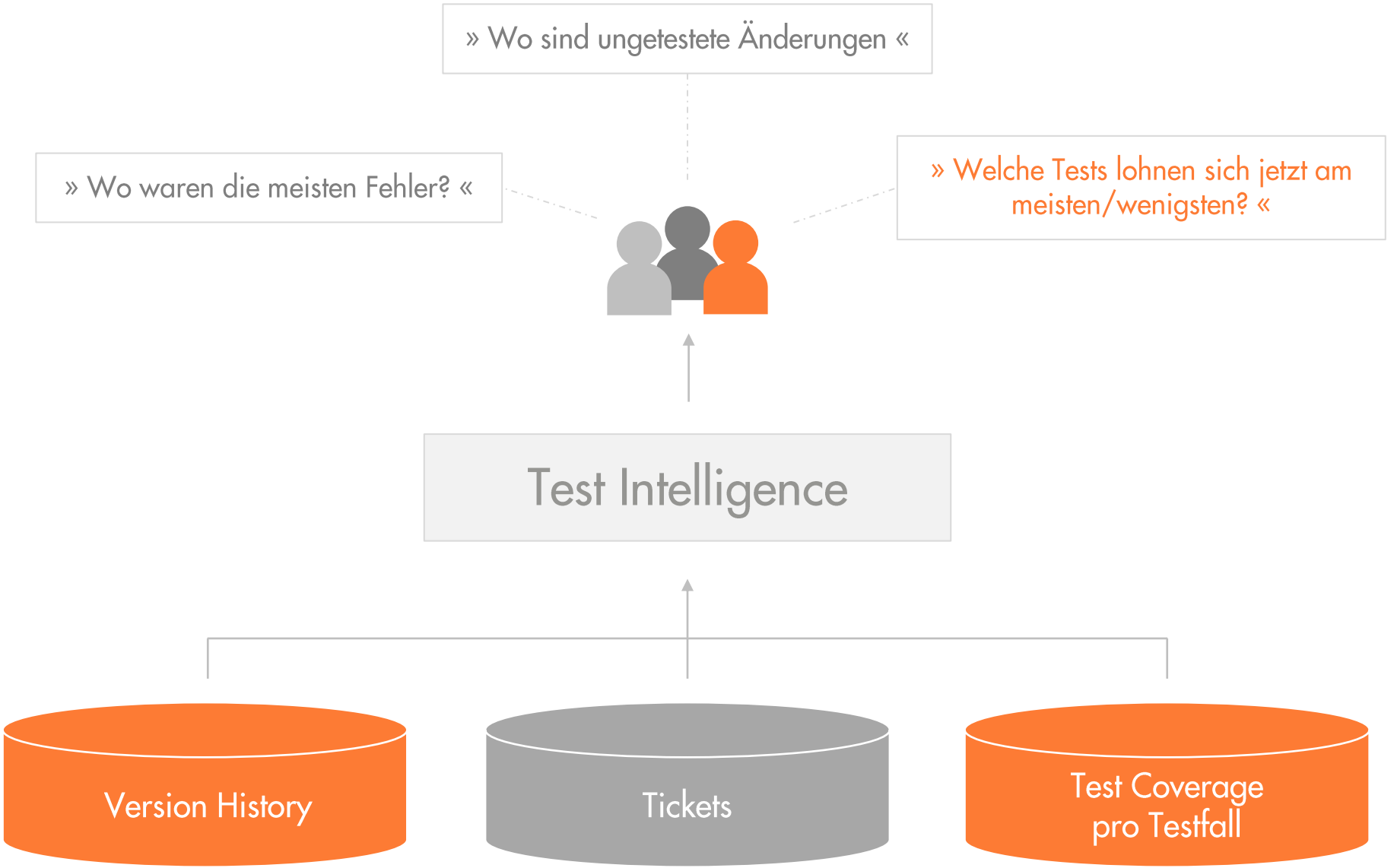
Vorher: nächtlich „Re-run all“

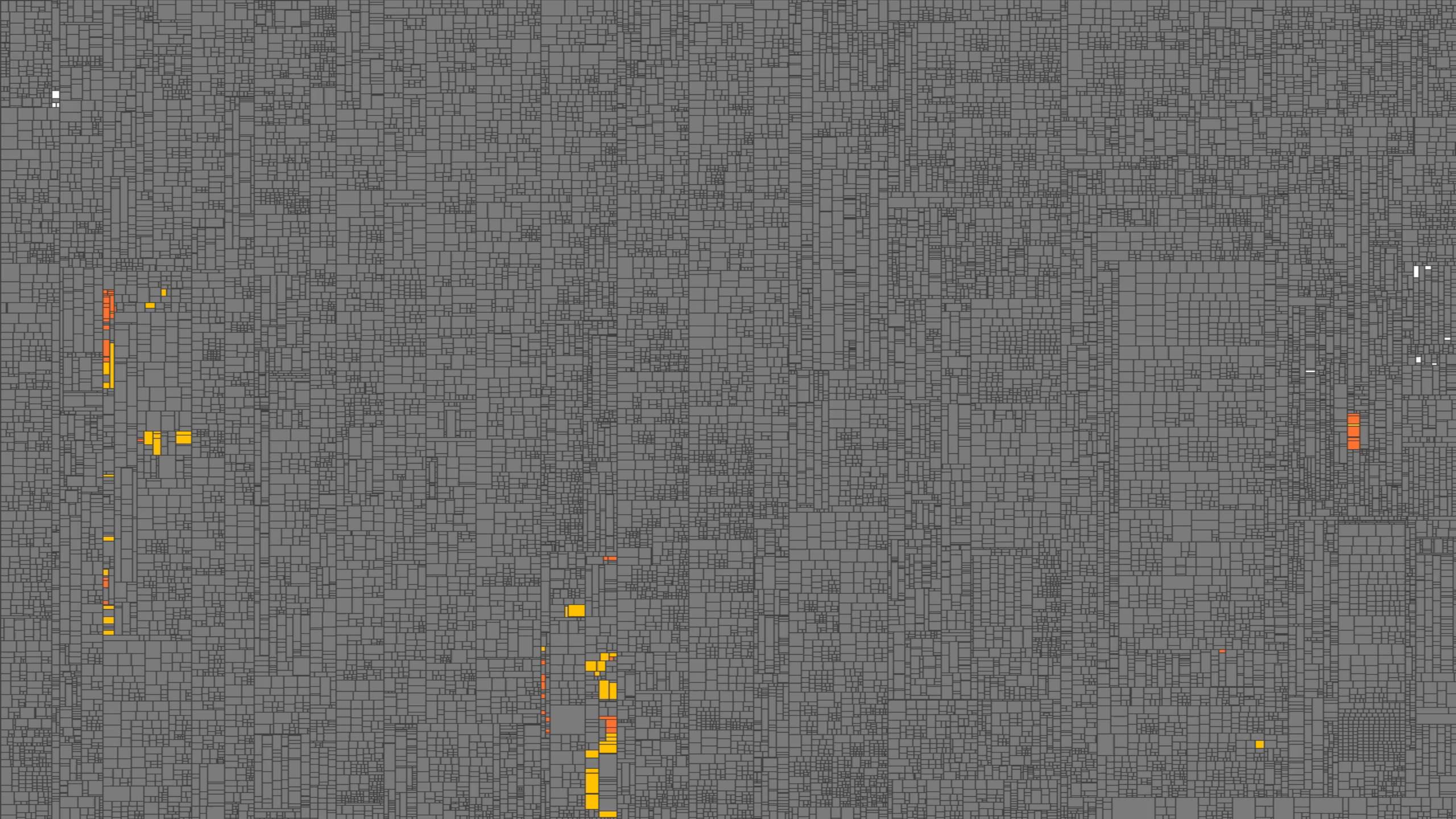
Feedback erst spät am nächsten Tag

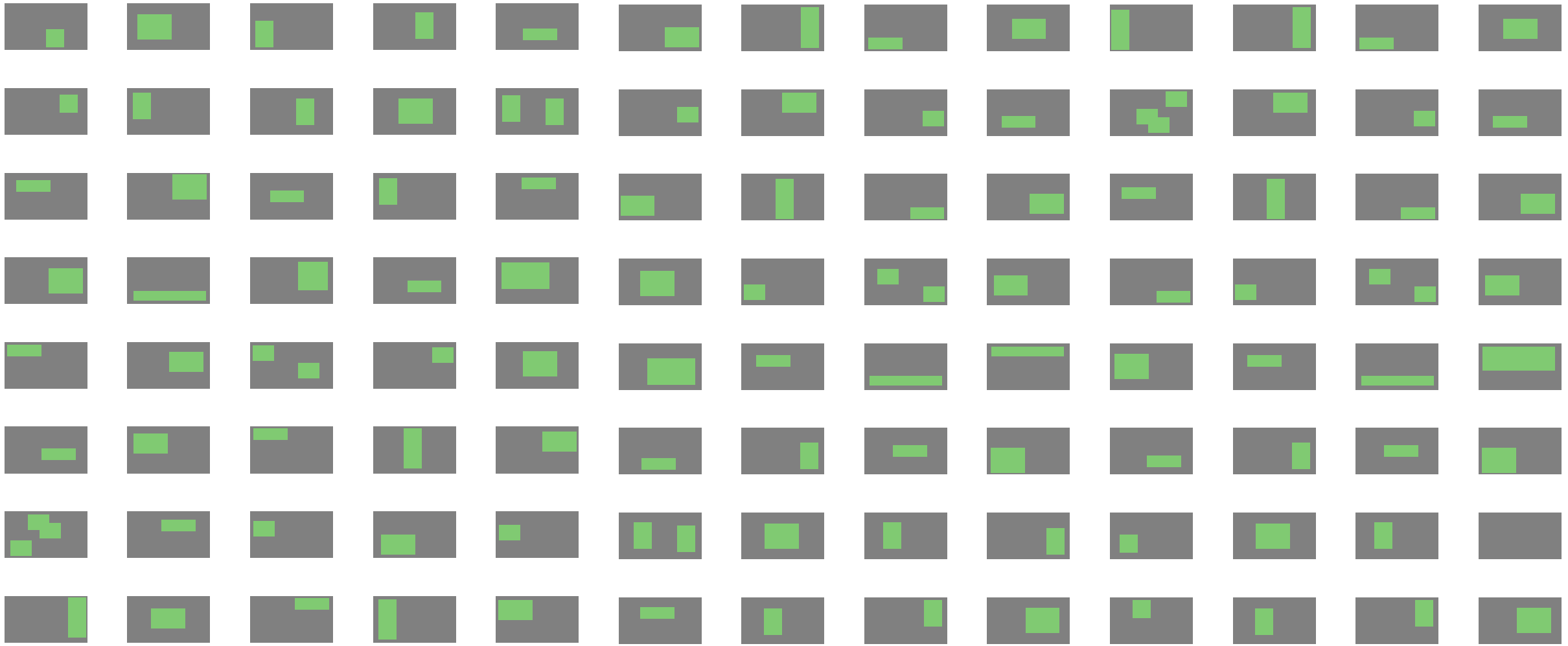
Tagsüber 1h Testbudget für Pareto-Testliste

Weiterhin: nächtlich „Re-run all“

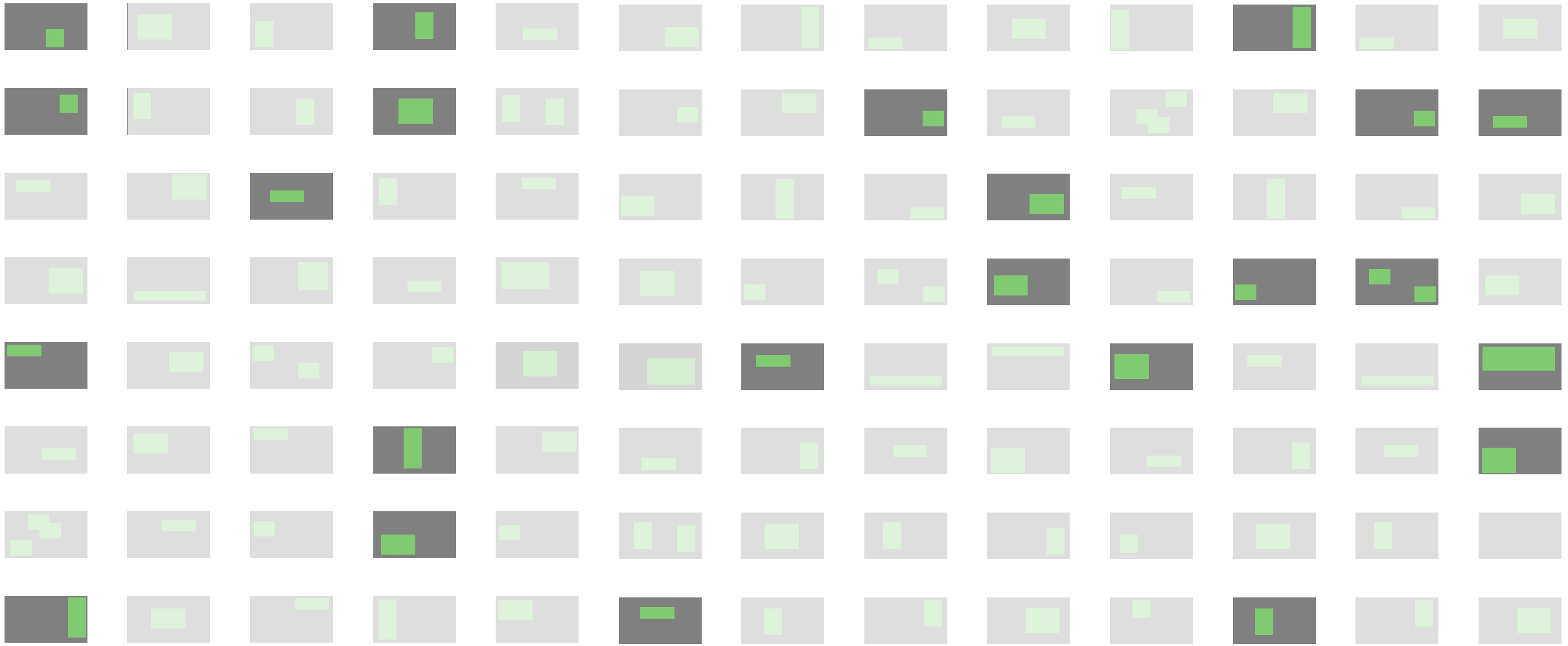
Mit nur 5% der Testlaufzeit finden wir **70% der fehlgeschlagenen Commits**


» Wo sind ungetestete Änderungen «
» Wo waren die meisten Fehler? «
» Welche Tests lohnen sich jetzt am meisten/wenigsten? «
Test Intelligence
Version History
Tickets
Test Coverage pro Testfall

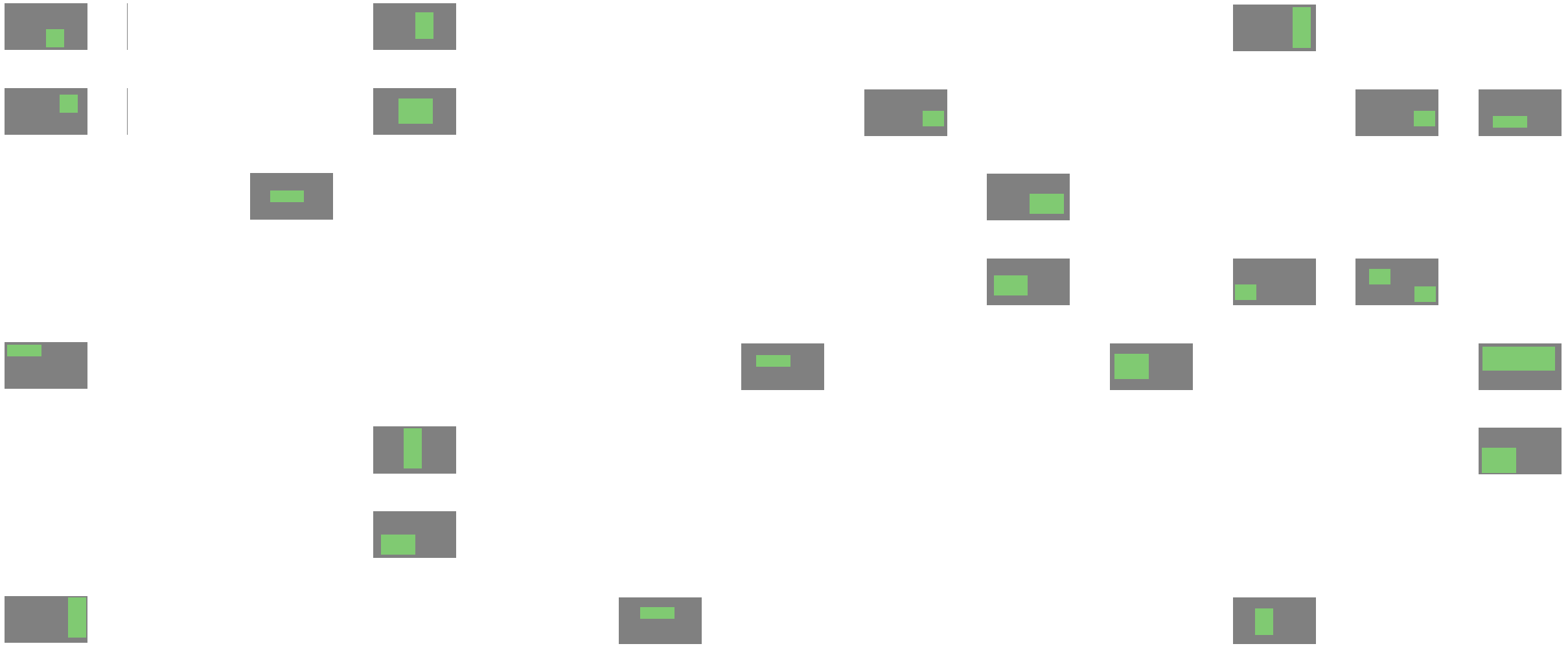




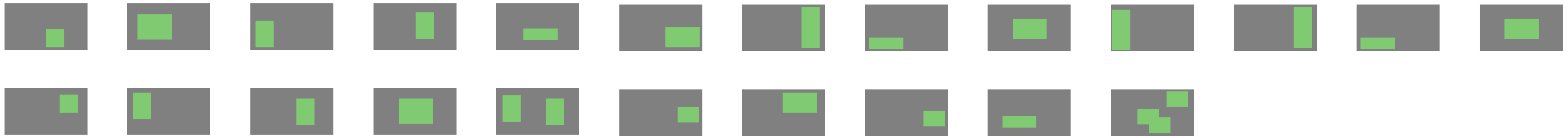
Schritt 1: Selektion betroffener Testfälle



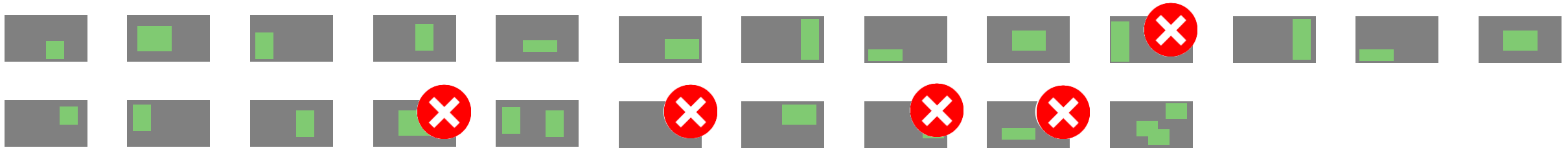
Schritt 1: Selektion betroffener Testfälle



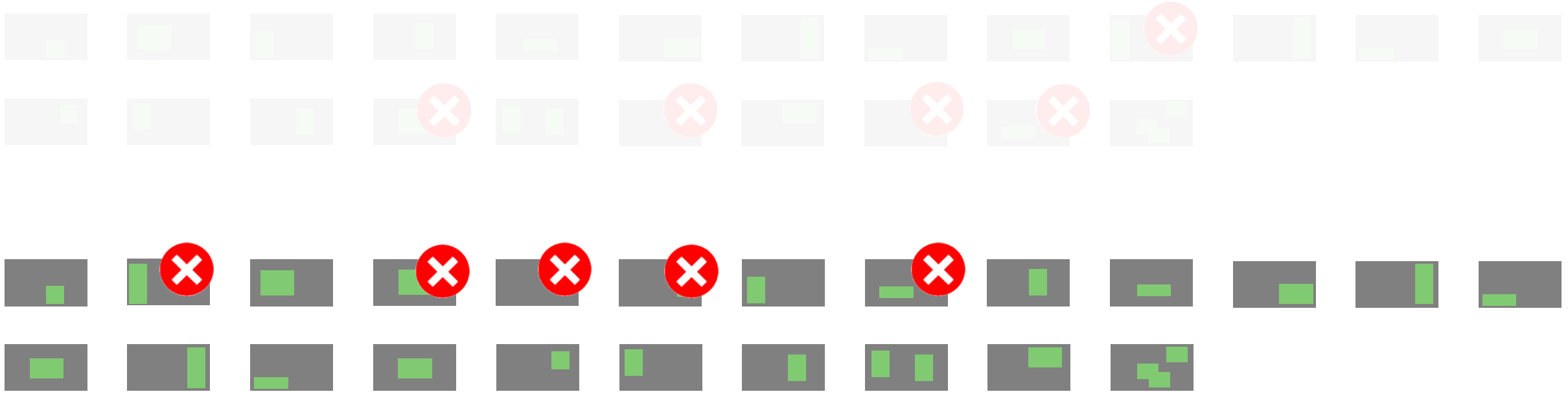
Schritt 1: Selektion betroffener Testfälle

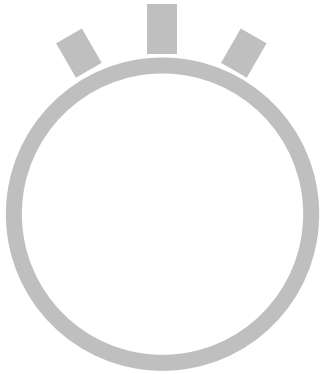


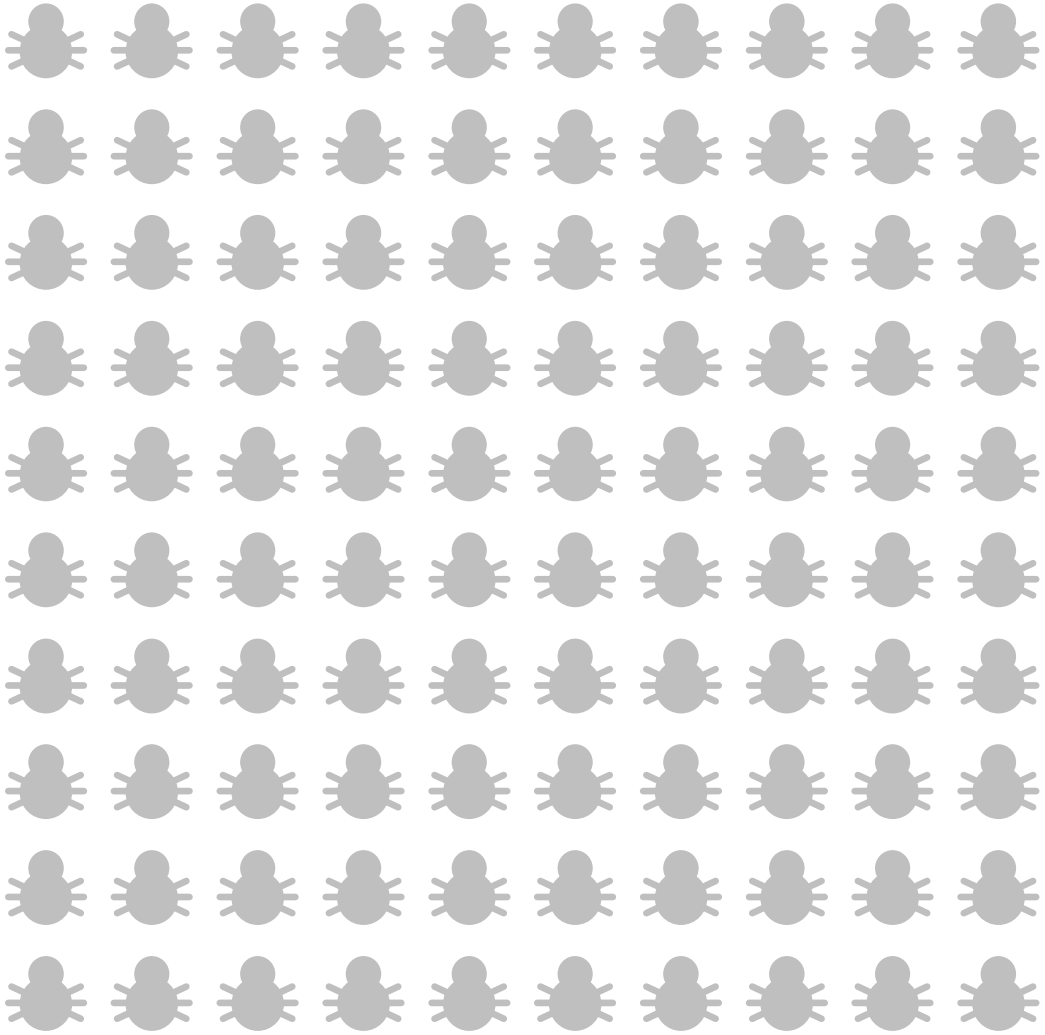
Schritt 2: Priorisierung selektierter Testfälle



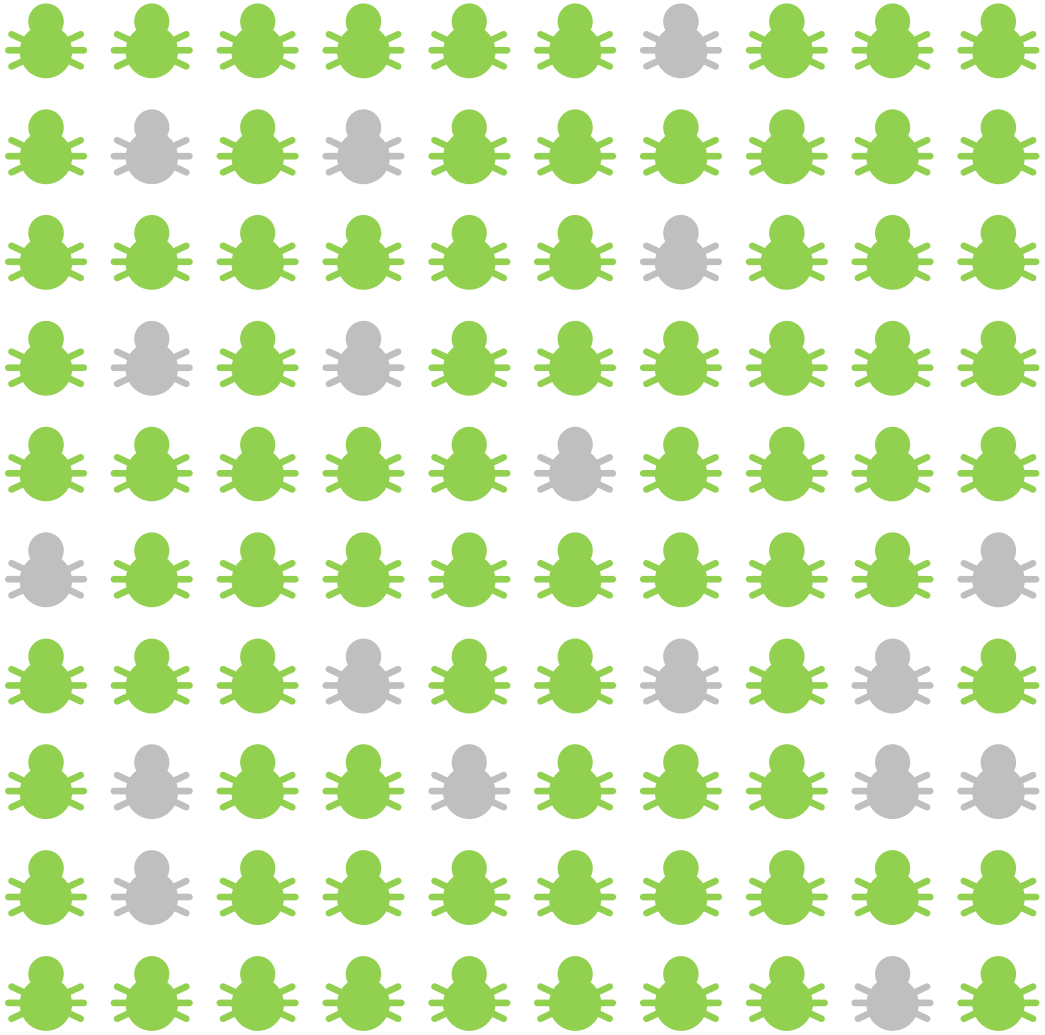
Schritt 2: Priorisierung selektierter Testfälle



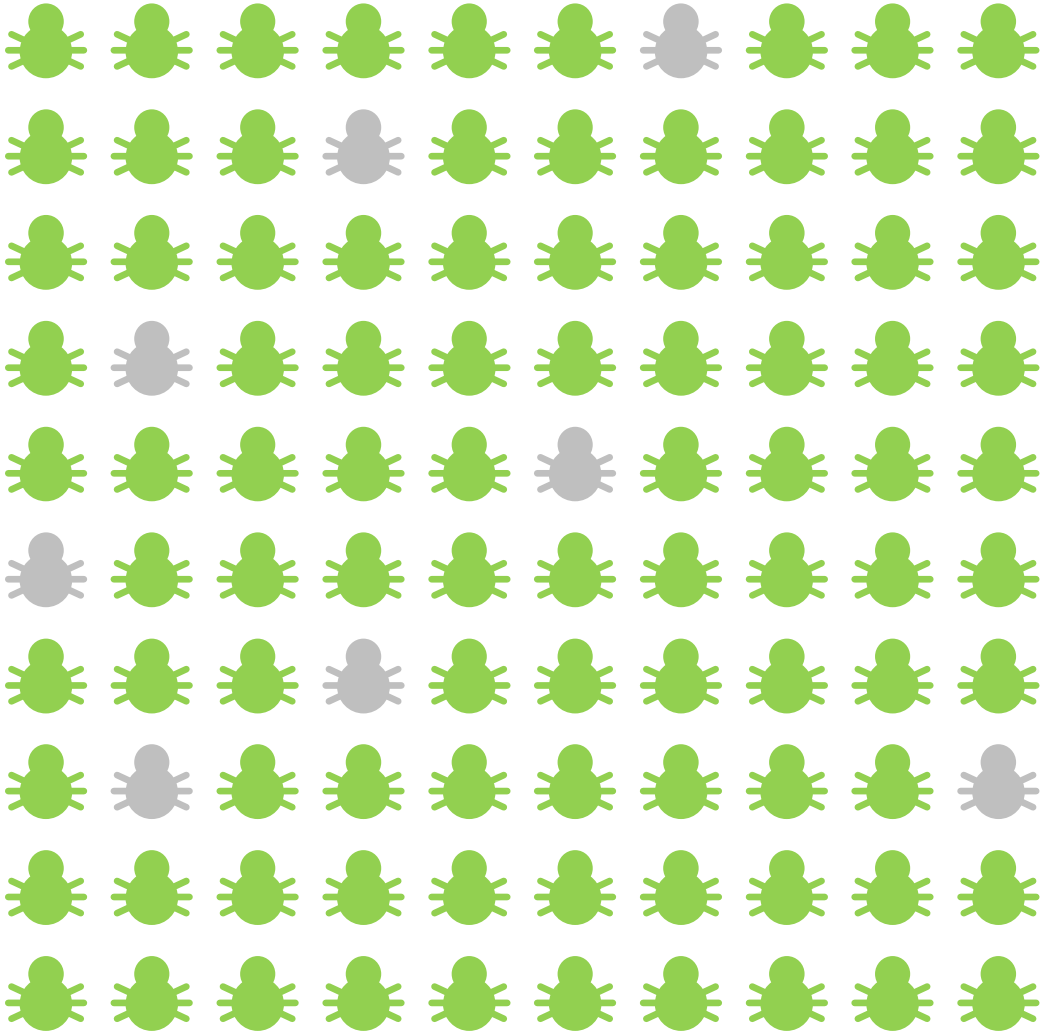




80 %



90 %



Test-Impact-Analyse bei Dolby Germany

Vor TIA

Testlaufzeiten von 3-12 Stunden durch eine Vielzahl von Parameterkombinationen.

Ziel

Test-Feedback aus der CI innerhalb von 10 Minuten.

Mit TIA

Seit einem Jahre kommt das Test-Feedback immer innerhalb von 10 Minuten, obwohl die Gesamtzahl der Tests im selben Zeitraum um 57% zunahm.

Test-Impact-Analyse

Tests werden für jeden Lauf **passend zu Änderungen** ausgewählt

90% der Fehler in 2% der Zeit

Erfordert kontinuierliche Messung der Coverage und Integration der Test-Auswahl in die CI / Testautomatisierung.

Stärkere Beschleunigung des Feedbacks (bei höherem Aufwand)

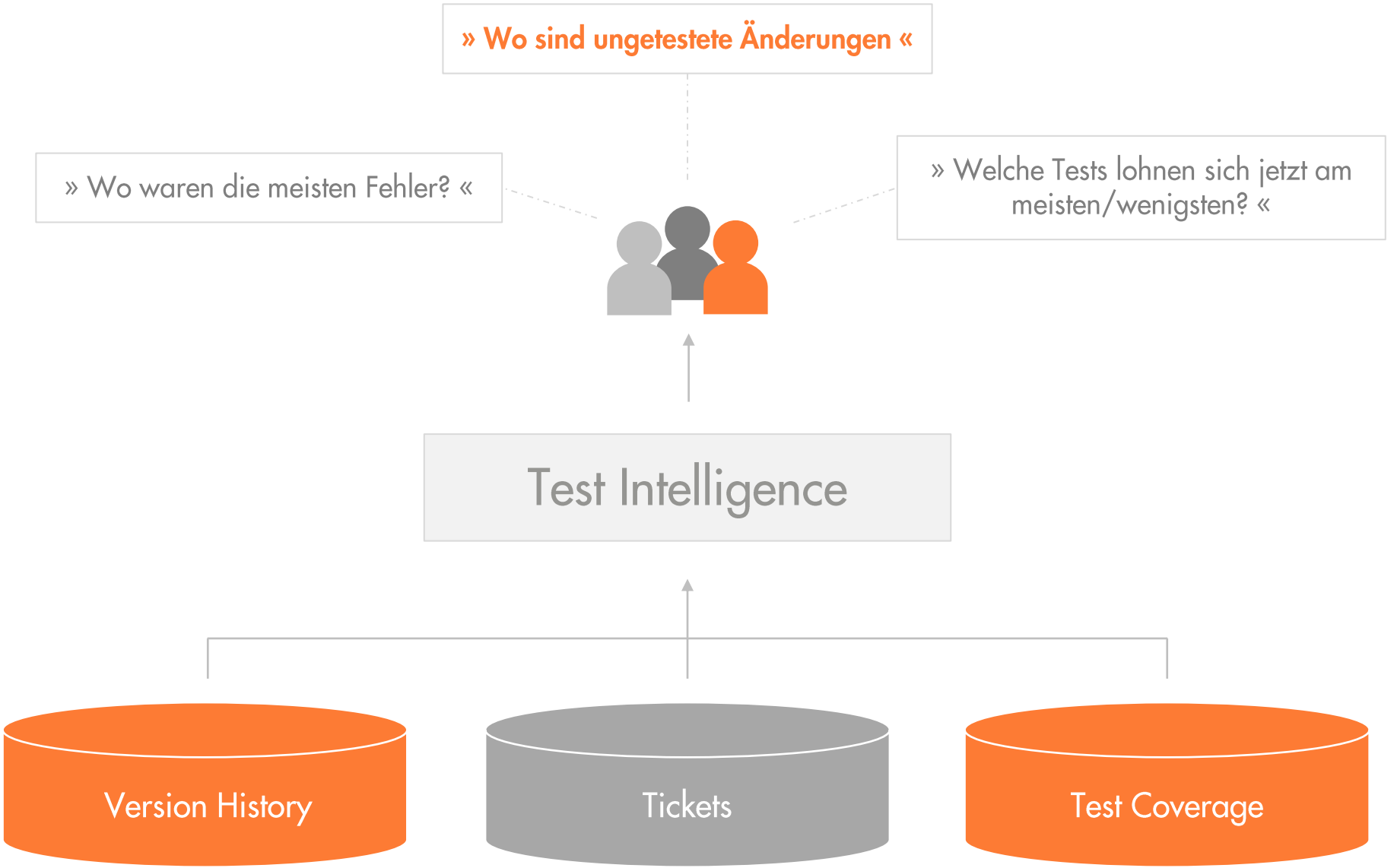
Pareto-Optimierung

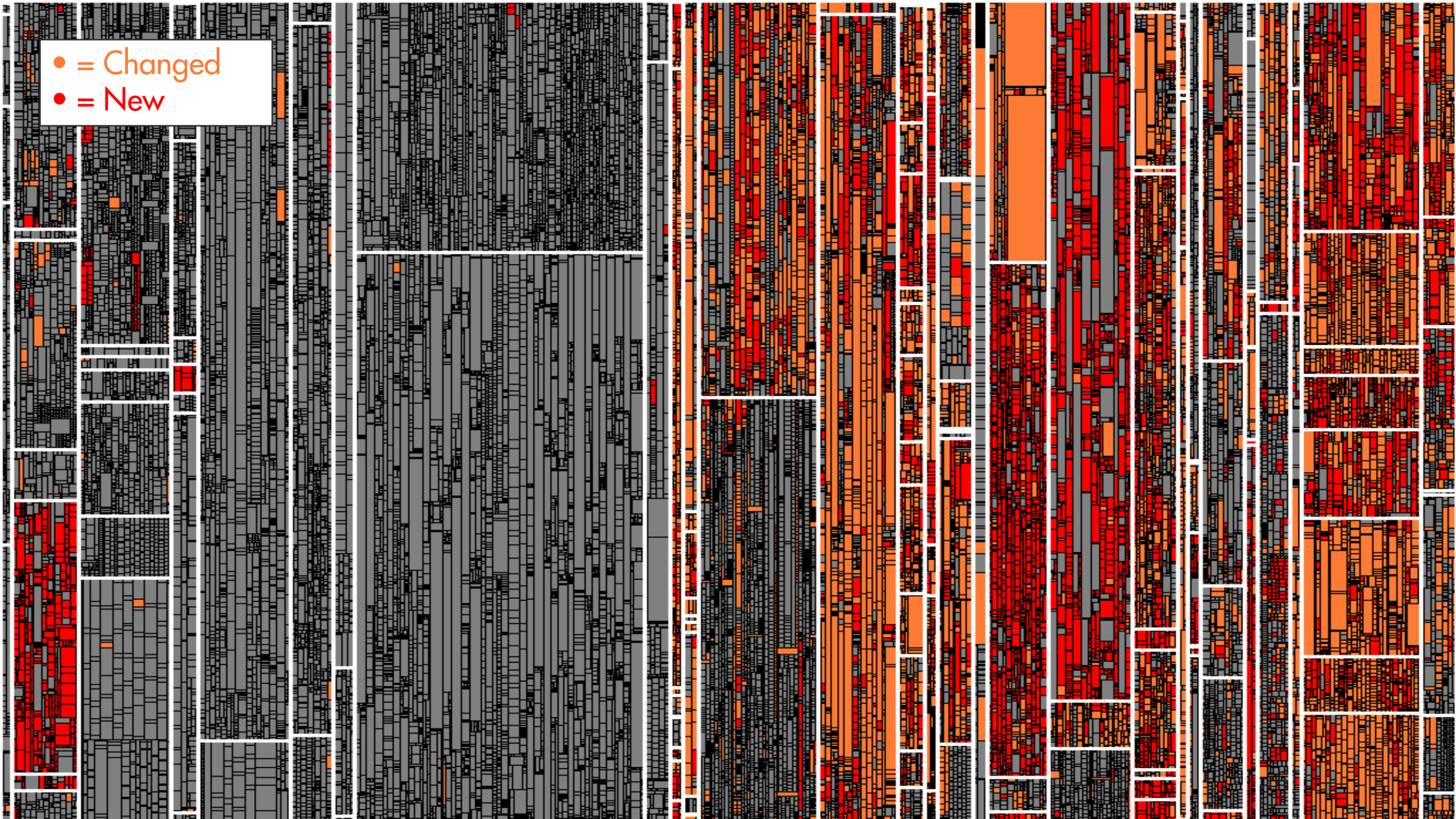
Tests werden **unabhängig von Änderungen** ausgewählt

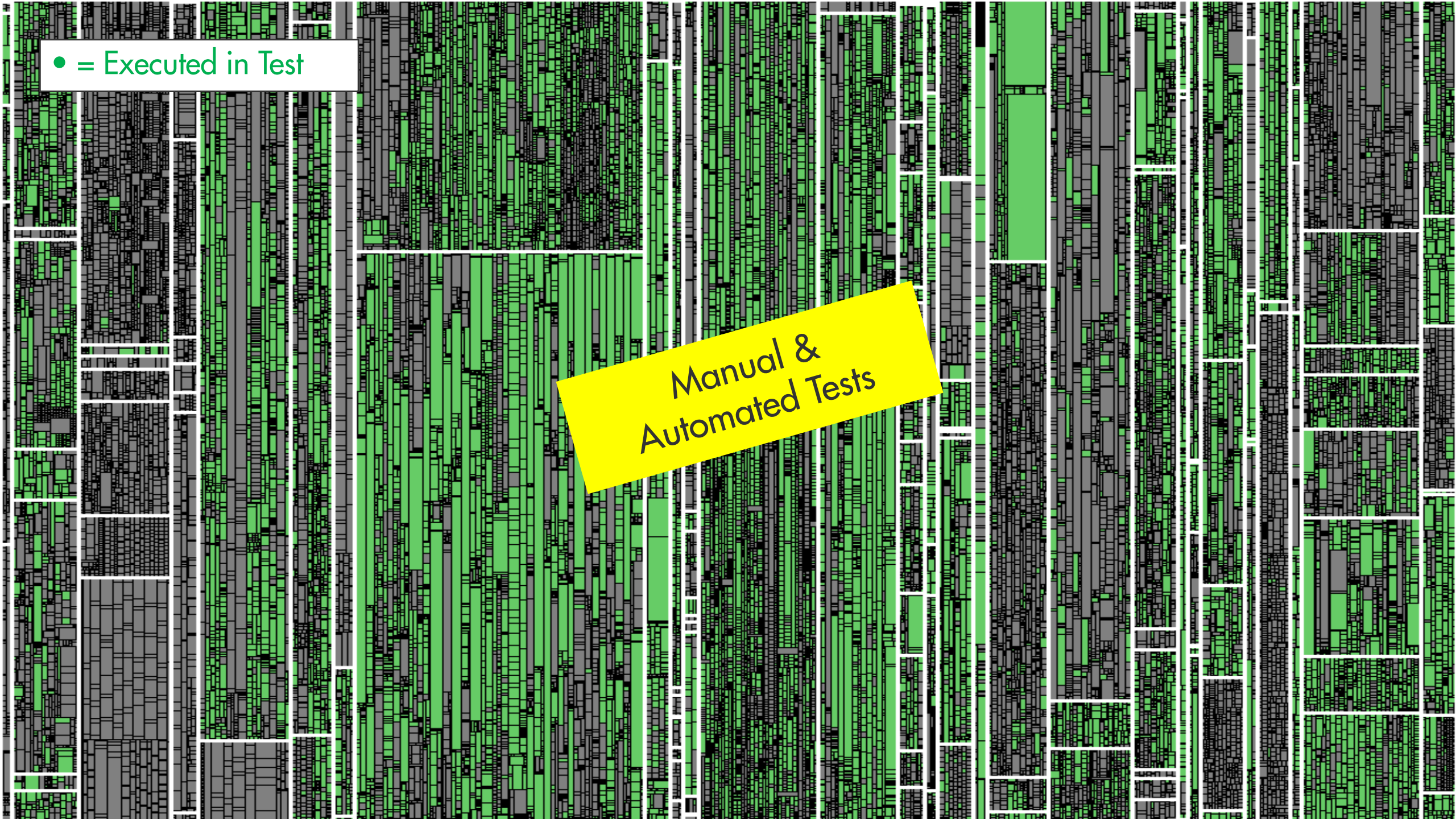
90% der Fehler in 11% der Zeit

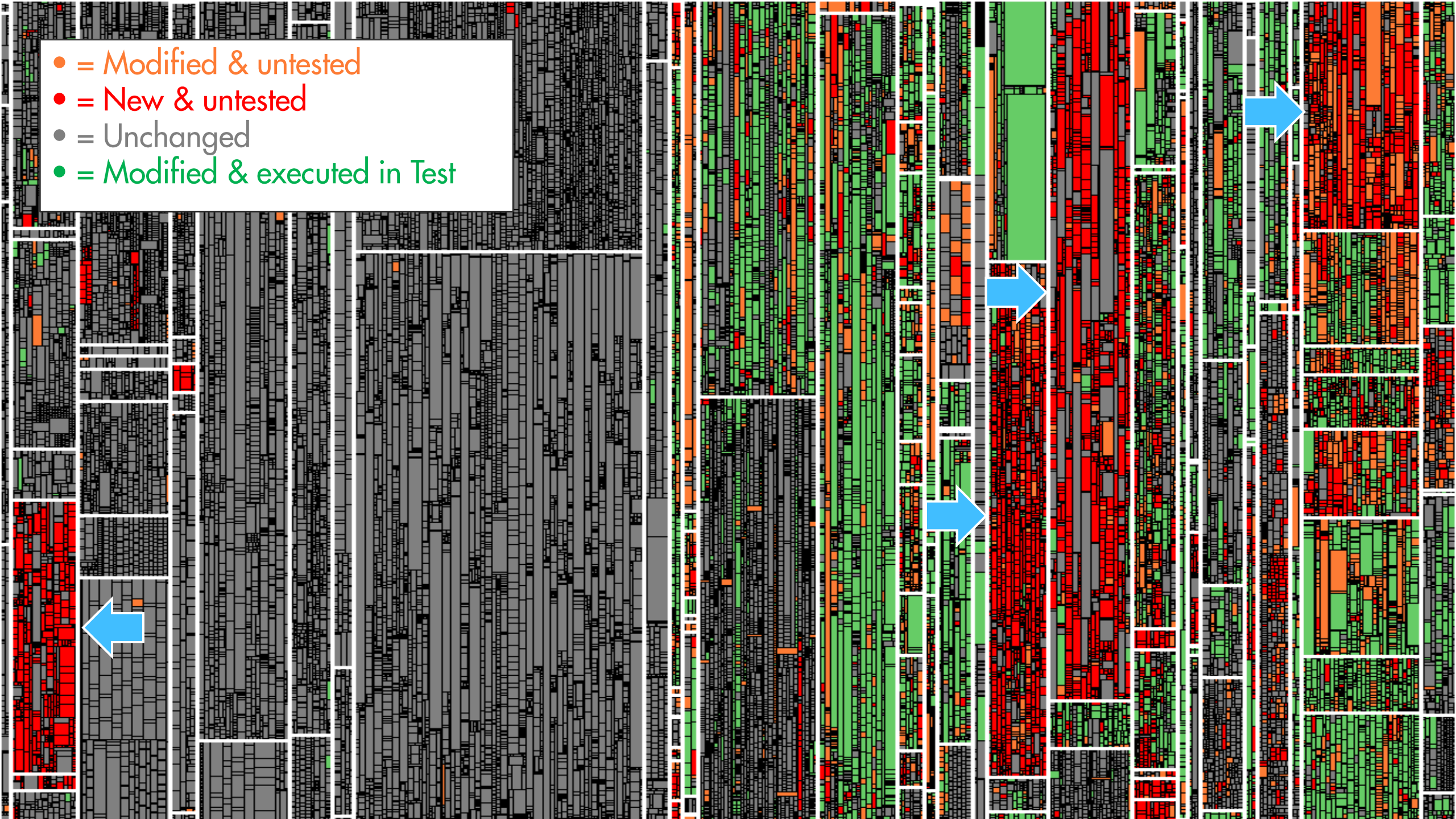
Einmalige Messung der Coverage reicht aus.

Viel geringerer Aufwand


» Wo sind ungetestete Änderungen «
» Wo waren die meisten Fehler? «
» Welche Tests lohnen sich jetzt am meisten/wenigsten? «
Test Intelligence
Version History
Tickets
Test Coverage


• = Changed
• = New


• = Executed in Test
Manual &
Automated Tests


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

Issue #	Subject			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

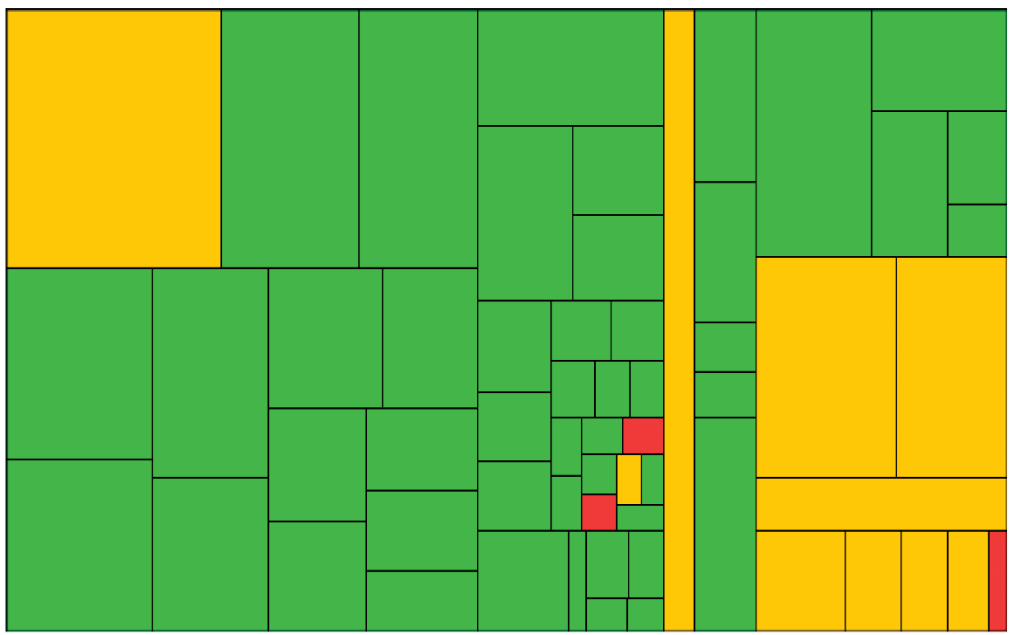
Assignee:

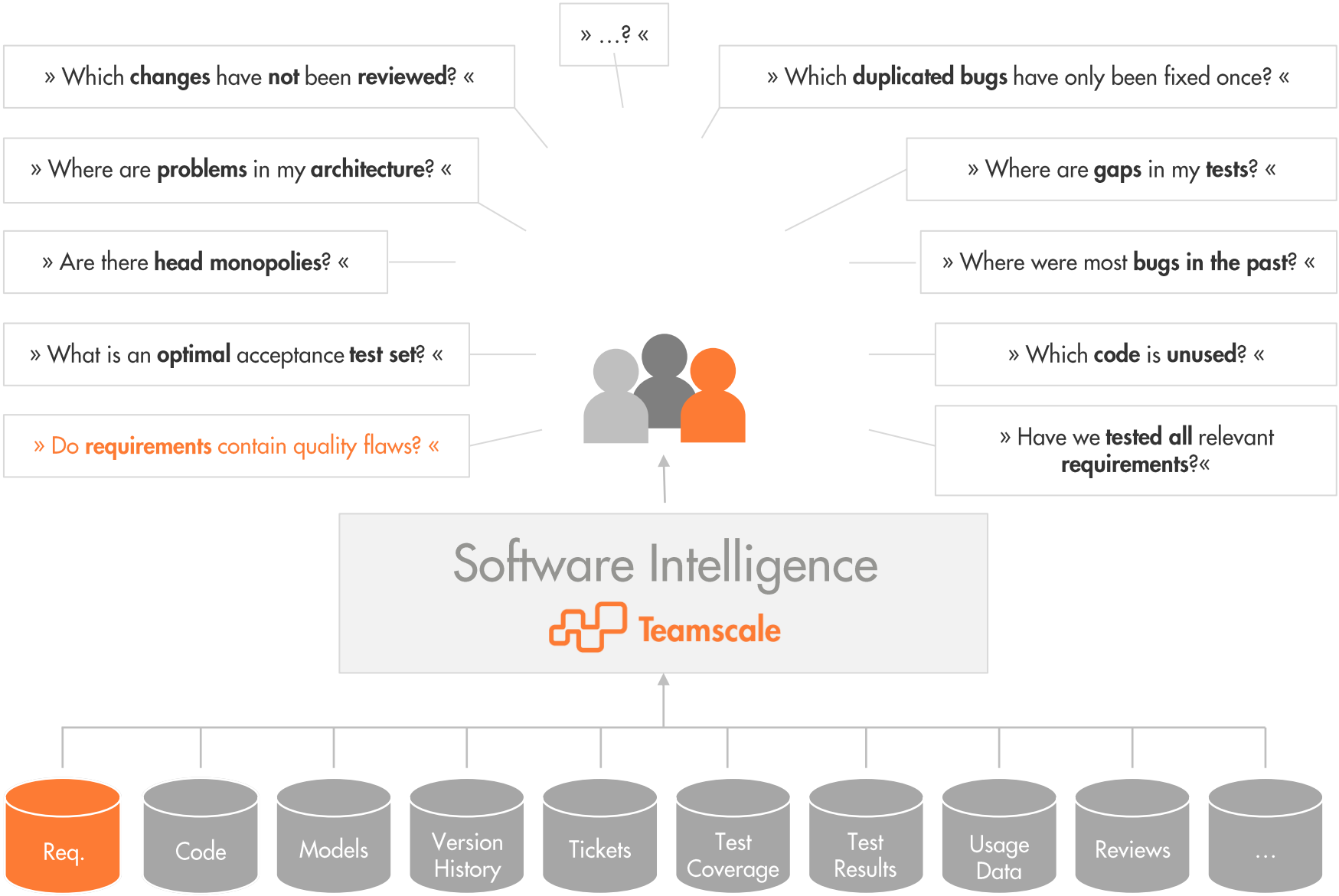
Project	Type	Priority	Resolution	Fix Version
TS	Maintenance	Normal	Green	Teamscale 4.5

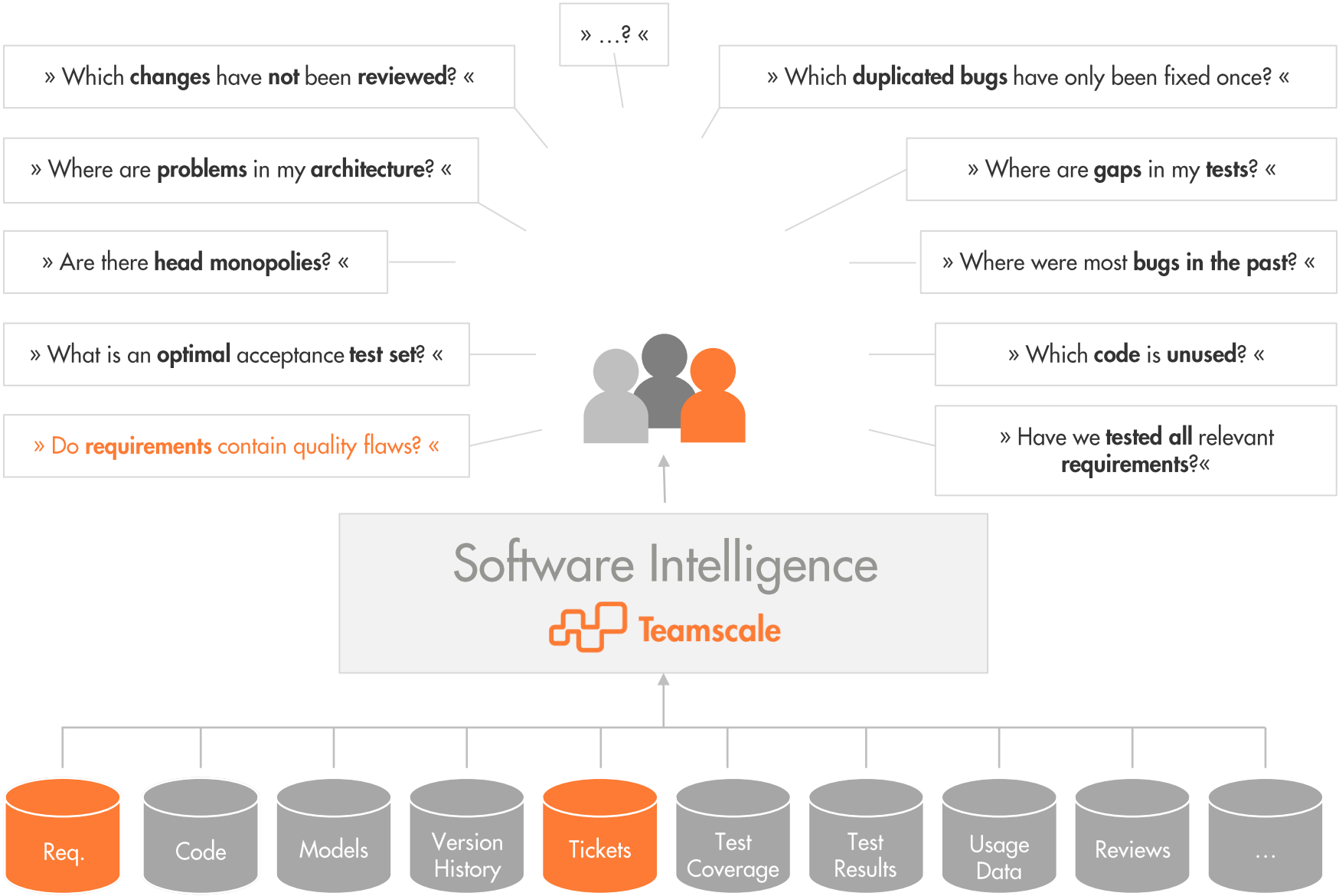
Component	Labels	Affected Version	Customer	Customer Issue
Backend	Performance			

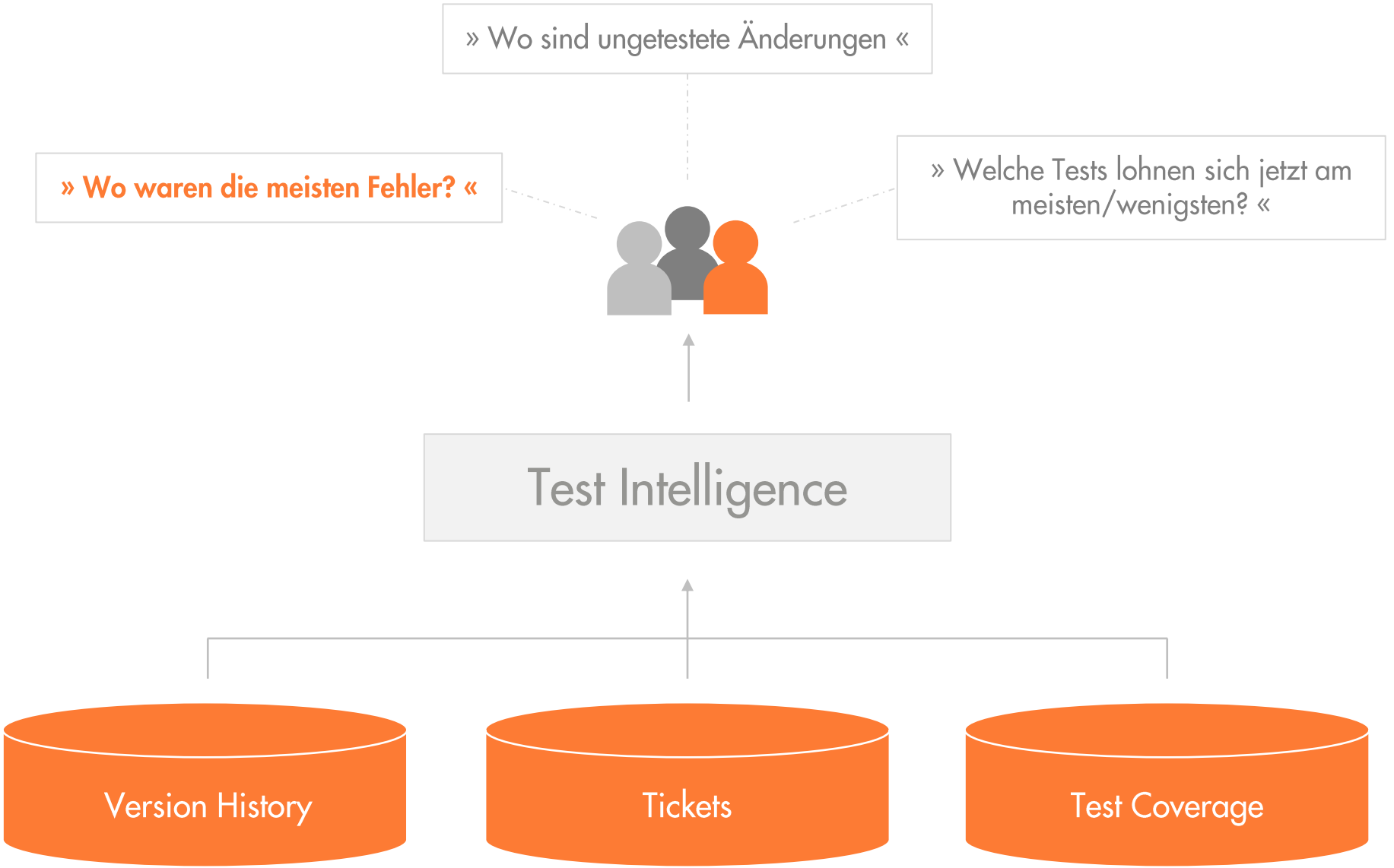
Epic Name	Freshdesk URL	Merge Request
		https://git.cqse.eu/cqse/teamscale/3621

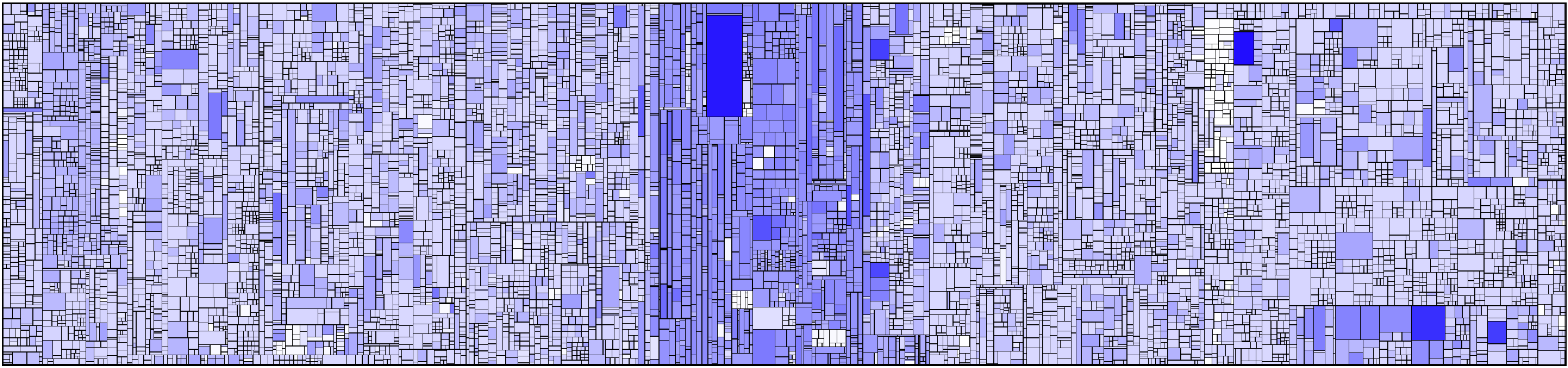
Aug 15 2018 12:37–Now | Test Gap: 22%

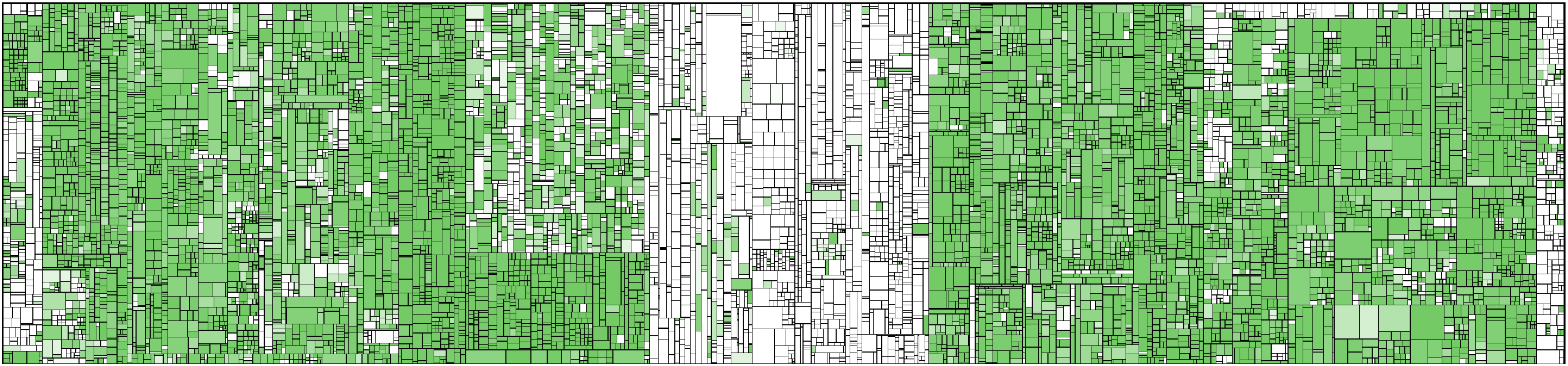


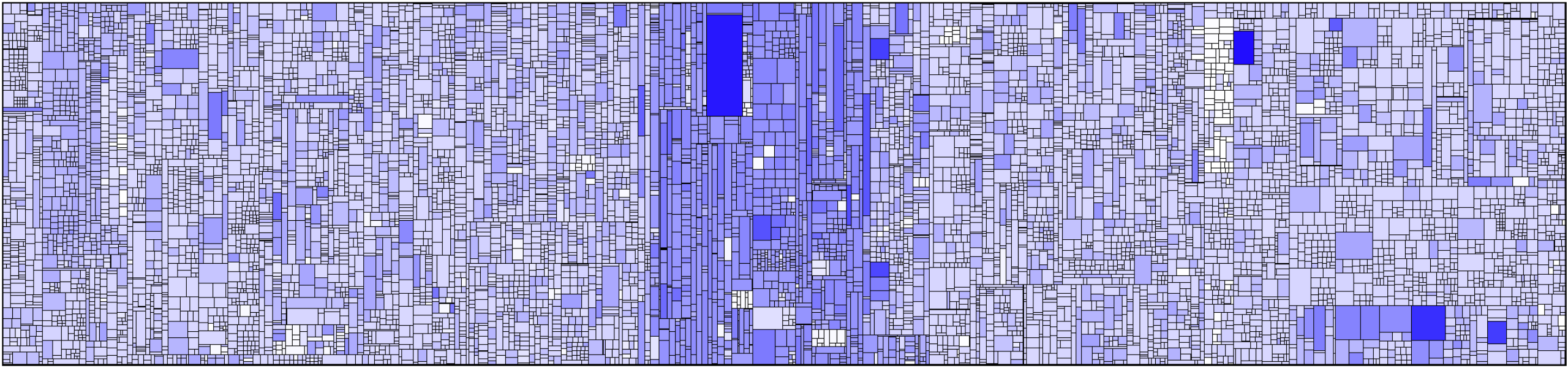

» …? «
» Which changes have not been reviewed? «
» Which duplicated bugs have only been fixed once? «
» Where are problems in my architecture? «
» Where are gaps in my tests? «
» Are there head monopolies? «
» Where were most bugs in the past? «
» What is an optimal acceptance test set? «
» Which code is unused? «
» Do requirements contain quality flaws? «
» Have we tested all relevant requirements?«
Software Intelligence
Teamscale
Req.
Code
Models
Version History
Tickets
Test Coverage
Test Results
Usage Data
Reviews
…

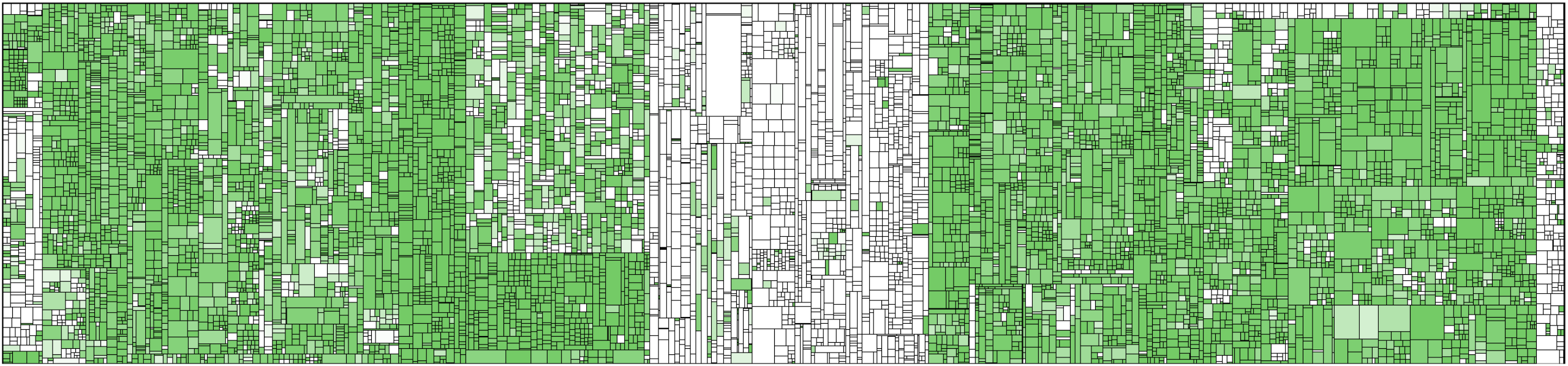

» …? «
» Which changes have not been reviewed? «
» Which duplicated bugs have only been fixed once? «
» Where are problems in my architecture? «
» Where are gaps in my tests? «
» Are there head monopolies? «
» Where were most bugs in the past? «
» What is an optimal acceptance test set? «
» Which code is unused? «
» Do requirements contain quality flaws? «
» Have we tested all relevant requirements?«
Software Intelligence
Teamscale
Req.
Code
Models
Version History
Tickets
Test Coverage
Test Results
Usage Data
Reviews
…


» Wo sind ungetestete Änderungen «
» Wo waren die meisten Fehler? «
» Welche Tests lohnen sich jetzt am meisten/wenigsten? «
Test Intelligence
Version History
Tickets
Test Coverage









CQSE Workshop

Test-Gap-Analyse

Ungetestete Änderungen im Quelltext aufdecken

04. Oktober
10:30-12:00 Uhr
cqse.eu/tga-2310-saa

Schnelles Feedback trotz langsamer Tests

Testselektion für historisch gewachsene Test-Suites

08. November
10:30-12:00 Uhr
cqse.eu/ts-2311-saa

Contact – Looking forward to discussions ☺

Dr. Elmar Jürgens · juergens@cqse.eu · +49 179 675 3863

Link zu Folien

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

CQSE