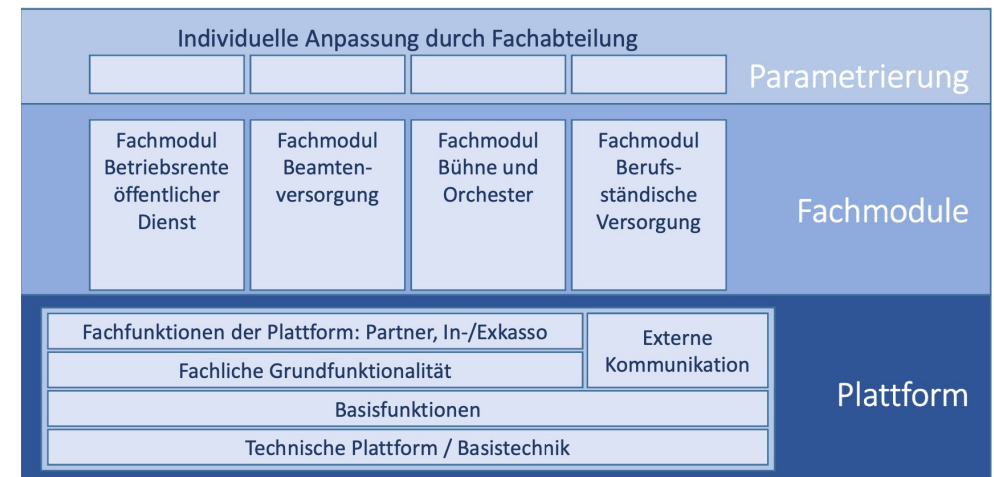


Schnelles Feedback von langlaufenden Tests

Erfahrungen aus der Zusammenarbeit mit der
Bayerischen Versorgungskammer (BVK)

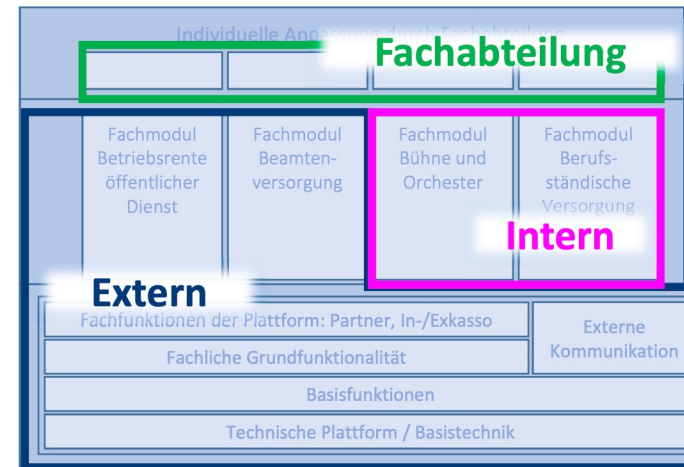
In-house Anwendung "Futura"

- Zentrale Businessprozesse (Mitgliederverwaltung, Rentenberechnung, ...)
- Seit 2007



In-house Anwendung "Futura"

- Zentrale Businessprozesse (Mitgliederverwaltung, Rentenberechnung, ...)
- Seit 2007
- Verteilte Entwicklung (intern / extern)
- Plattform + Fachmodule = ~1M SLOC Java Code
- Tests
 - Unit Tests und funktionale Tests
 - Integrationstests (GeVo-Tests, Testketten)
 - Manuelle Tests



Automatisierte Integrationstests

- Nächtliche Testläufe
- Durchschnittliche Laufzeit: 12h (bis 26h!)

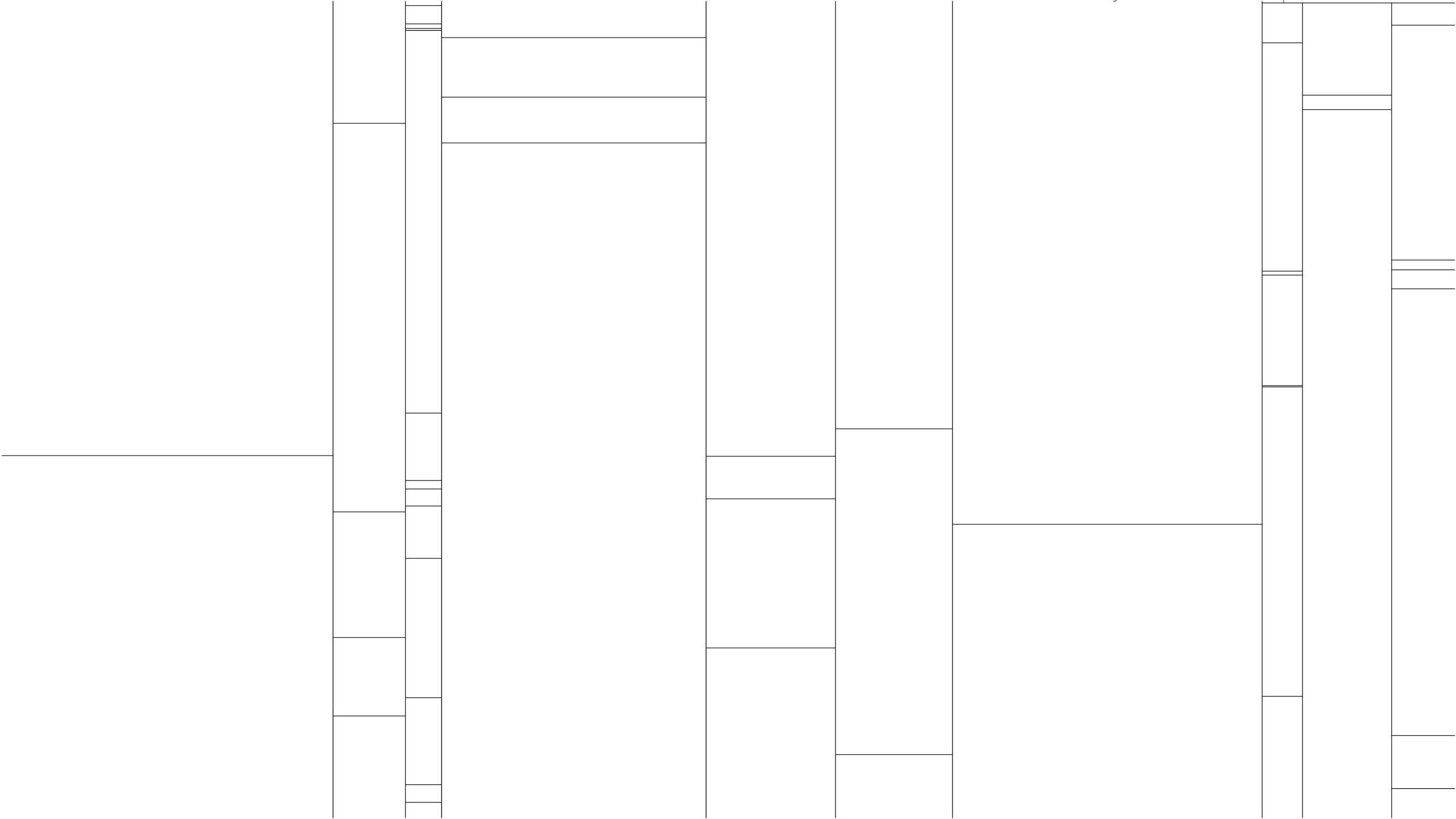
Typ	Anzahl	Laufzeit	*
Einzel-GeVo	3060	3h 41min	
Testketten	4095	6h 34min	
Batch	1860	1h 54min	
Summe	9015	12h 09min	
Unit + Functional Tests	6994	< 2min	

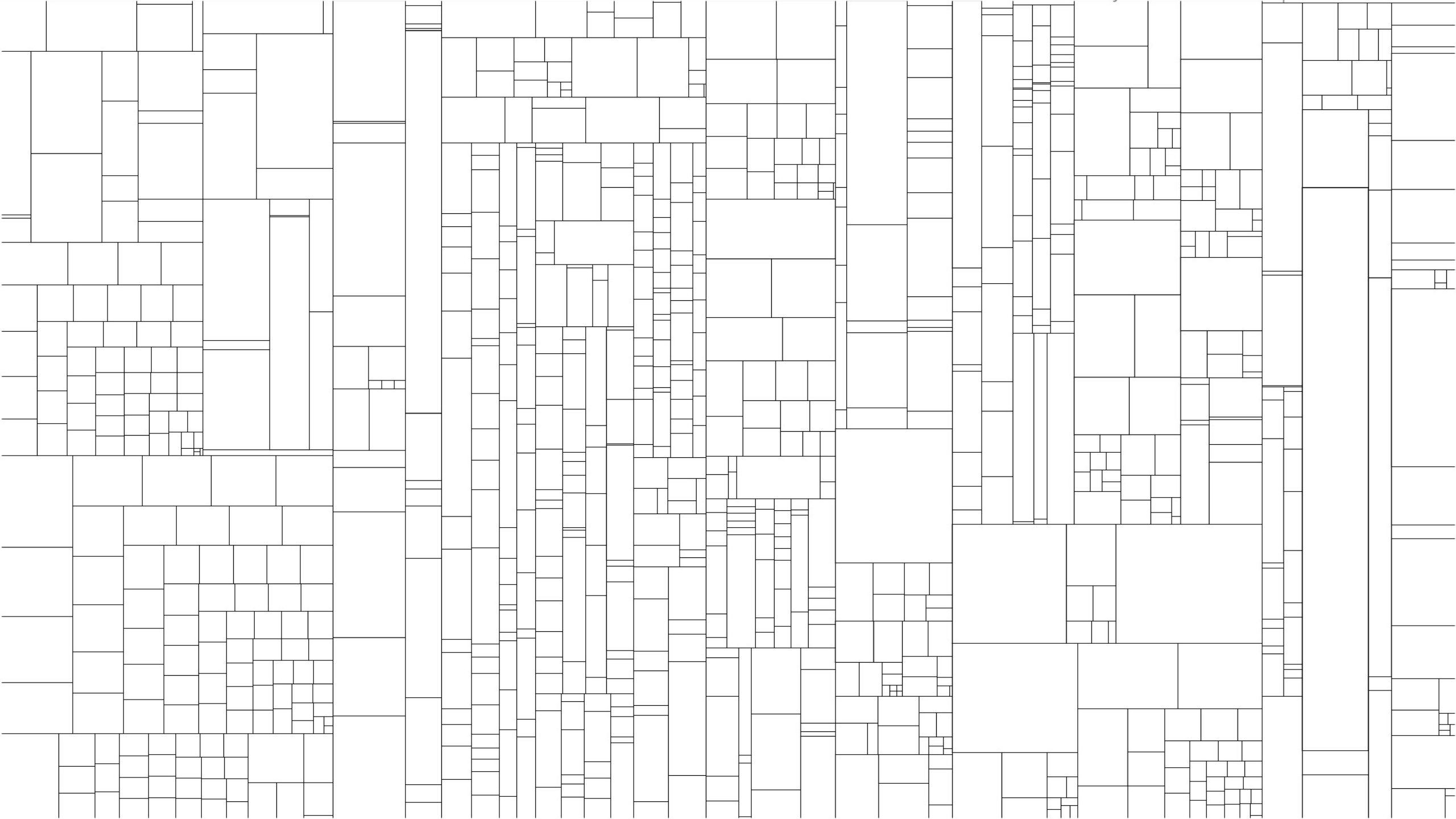
*Beispieldaten für Fachmodul „Bühne und Orchester“

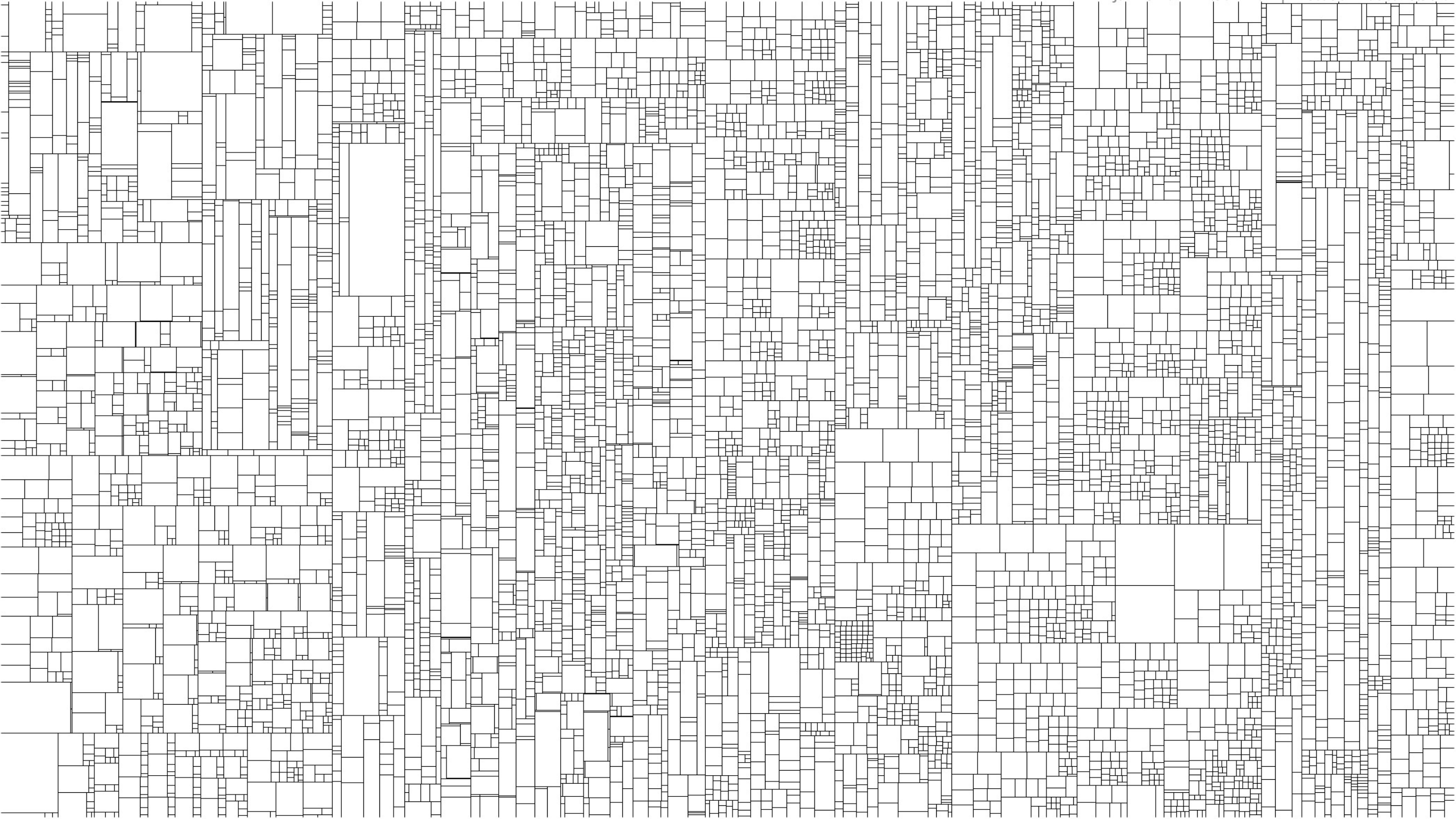
(Stand 03.05.2021)

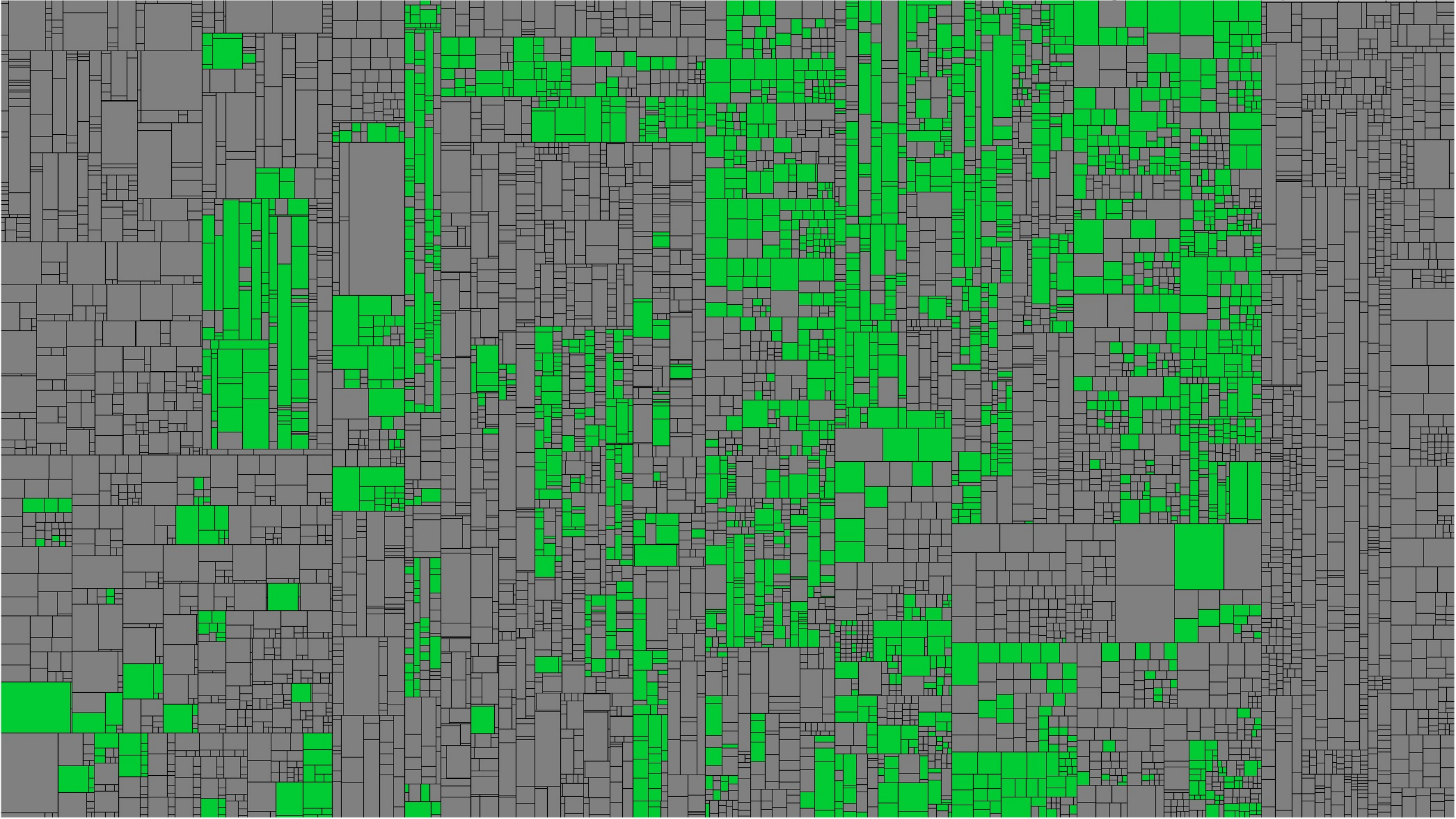
Können wir die Testlaufzeit reduzieren?

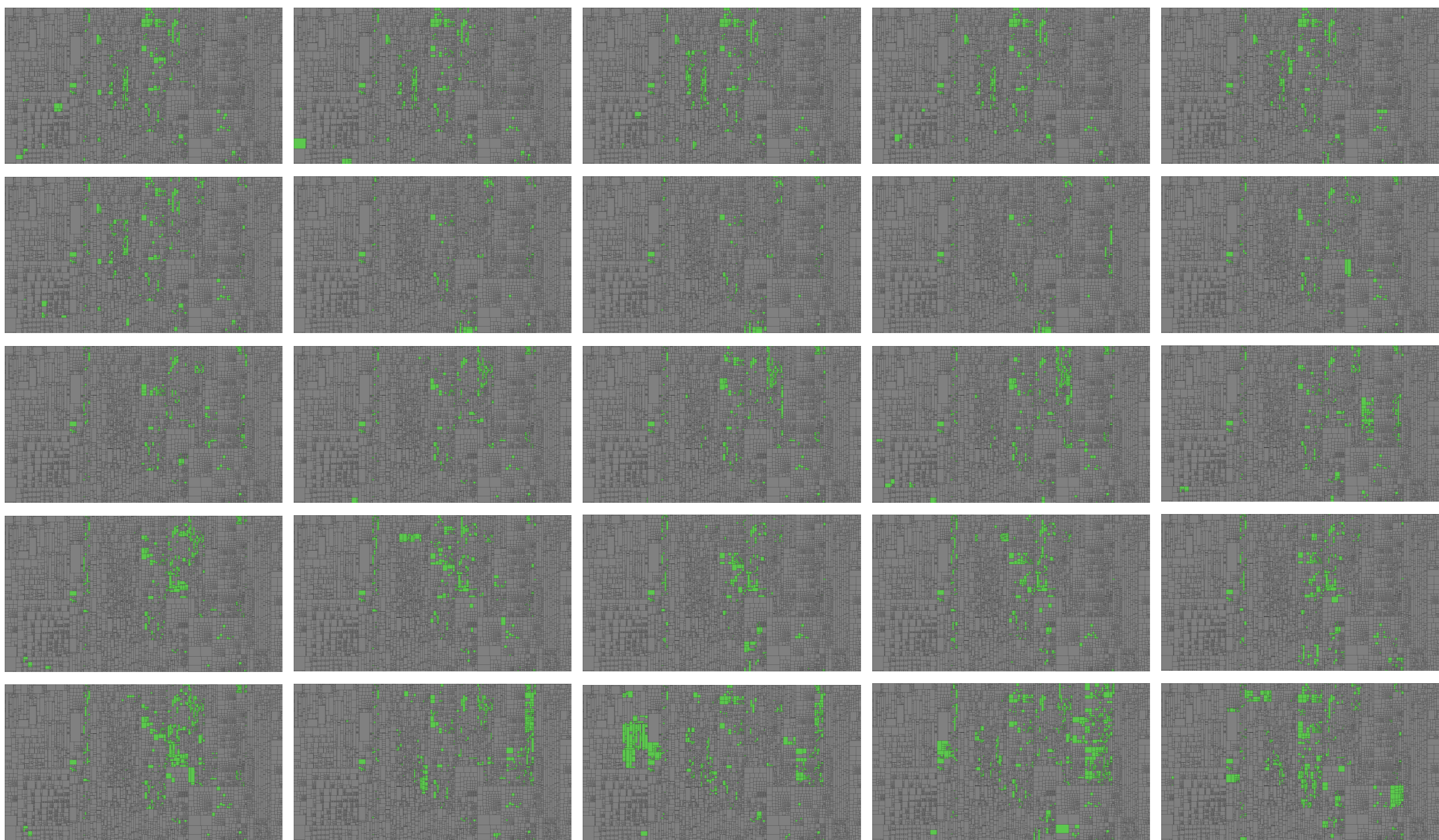
Brauchen wir immer alle Tests um Änderungen abzudecken?



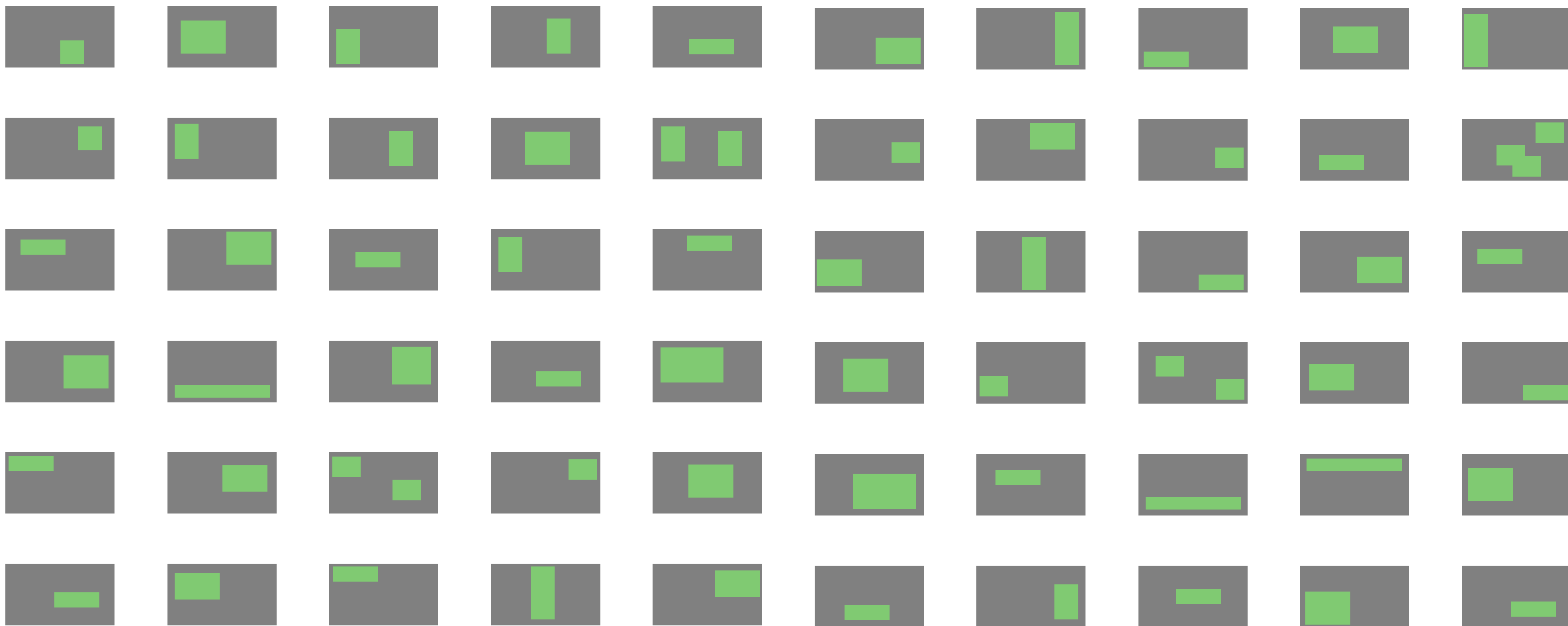


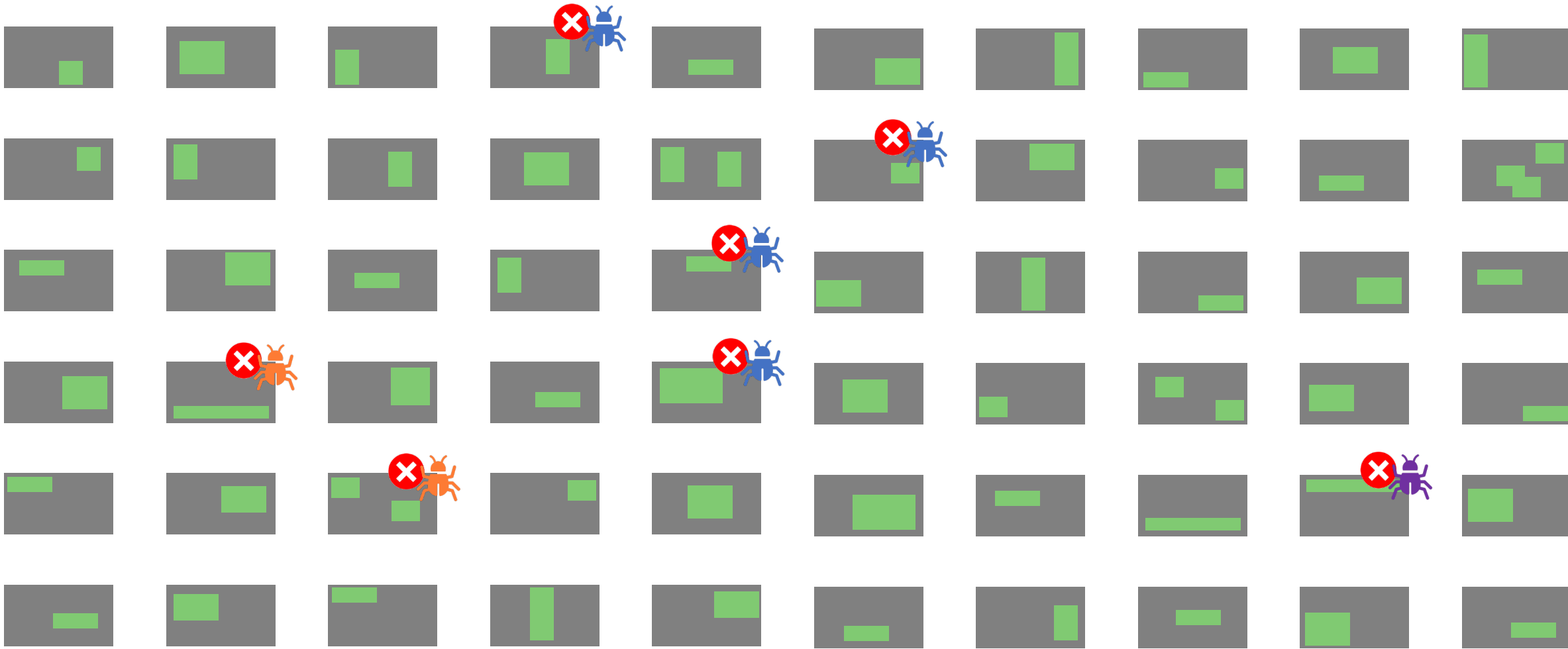




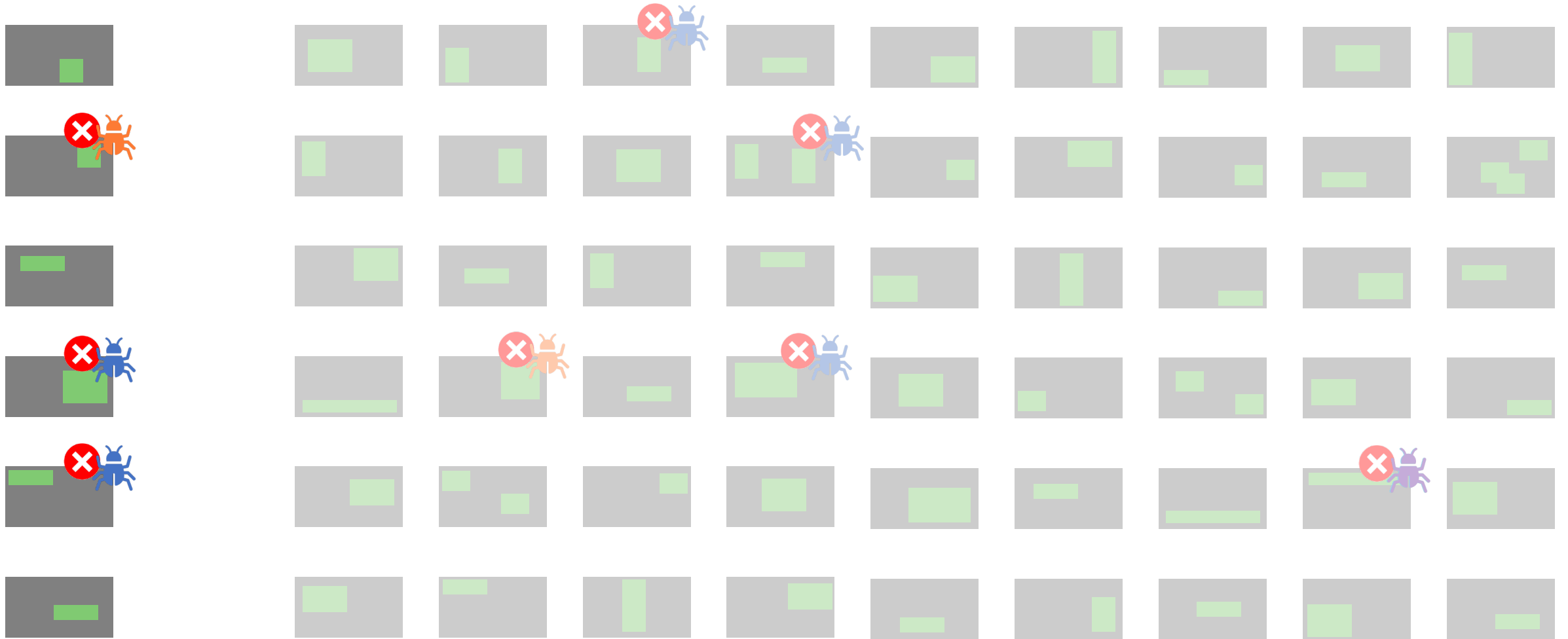


Wie entscheiden wir, welche Tests ausgeführt werden?

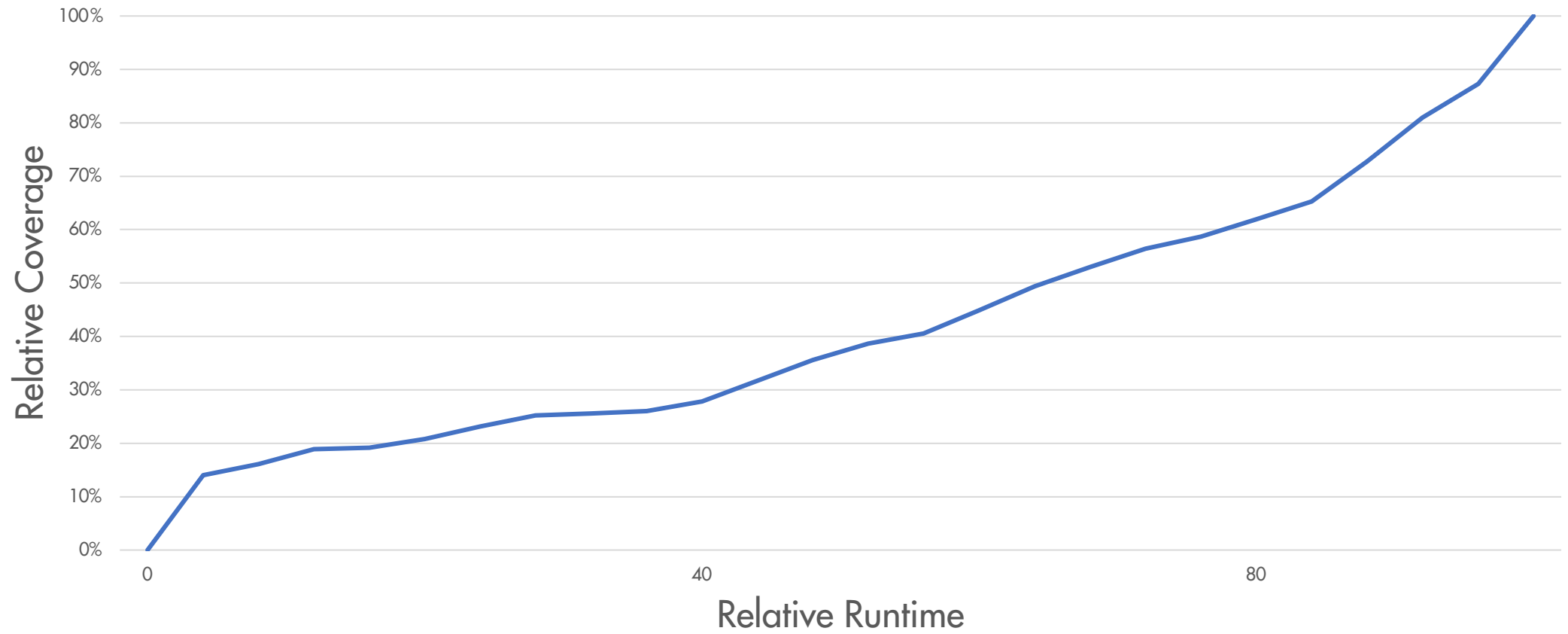




Pareto-Testing



Time vs Code Coverage



Testfall-spezifische
Coverage

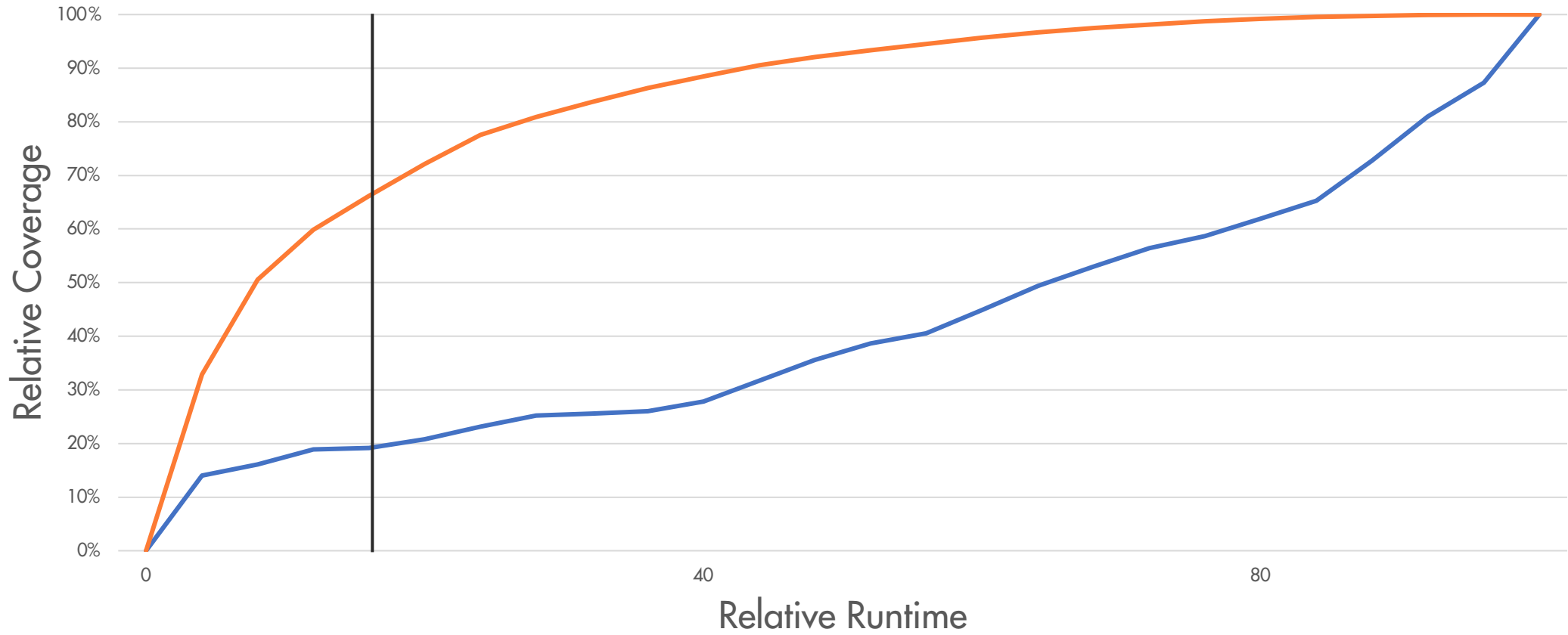


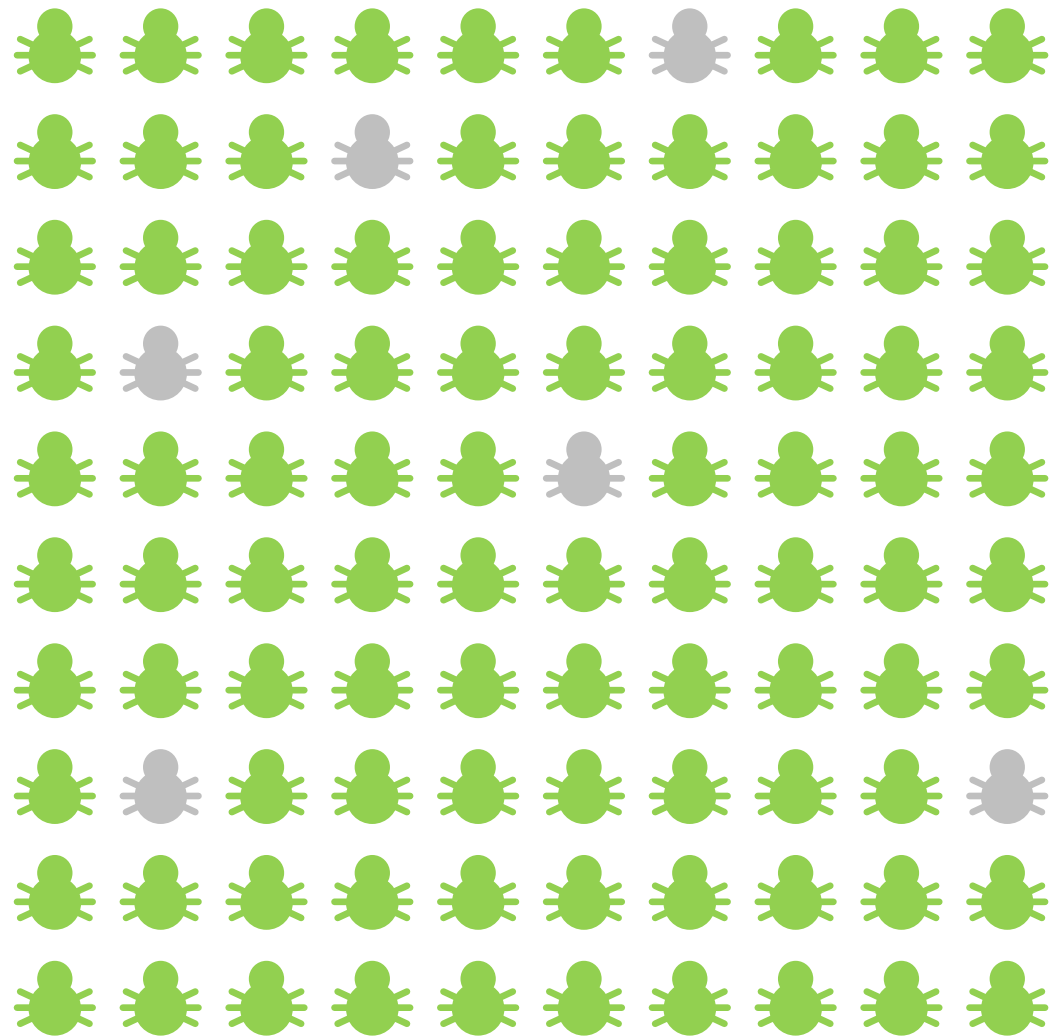
Pareto-Testing



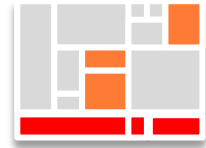
Selektierte & priorisierte
Testfälle

Time vs Code Coverage





Änderungen



Test-Impact-Analyse

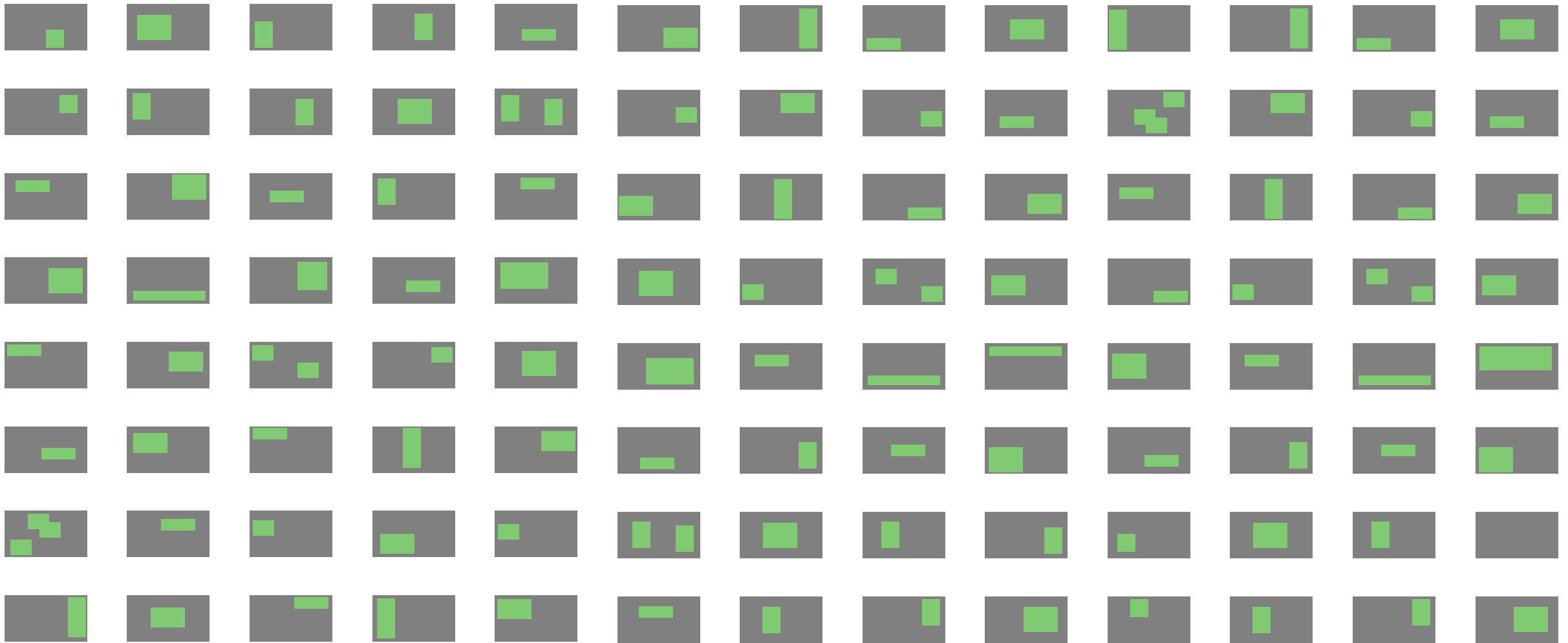


Selektierte & priorisierte
Testfälle

Testfall-spezifische
Coverage



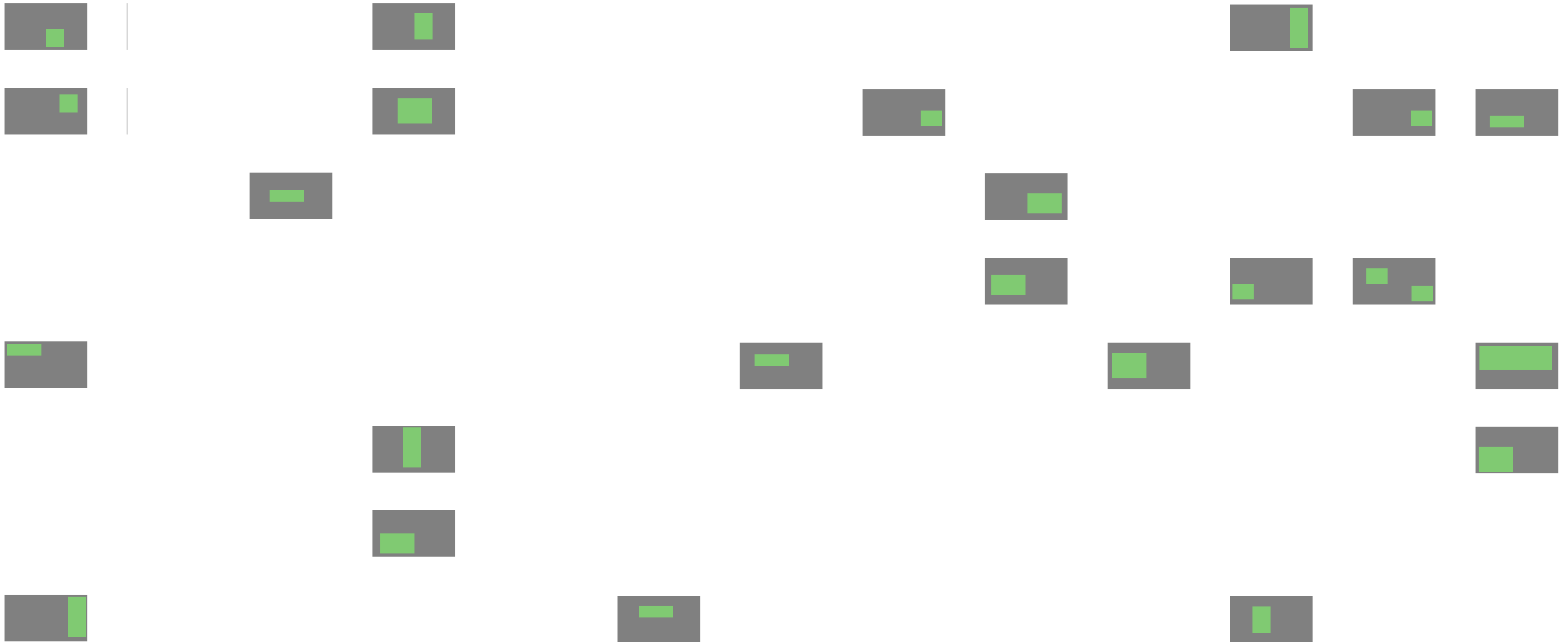
Test-Impact-Analyse



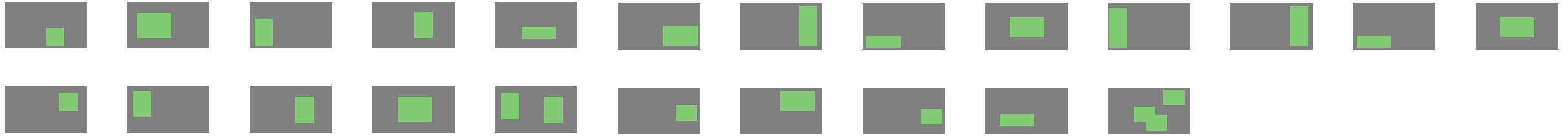
Schritt 1: Selektion betroffener Testfälle



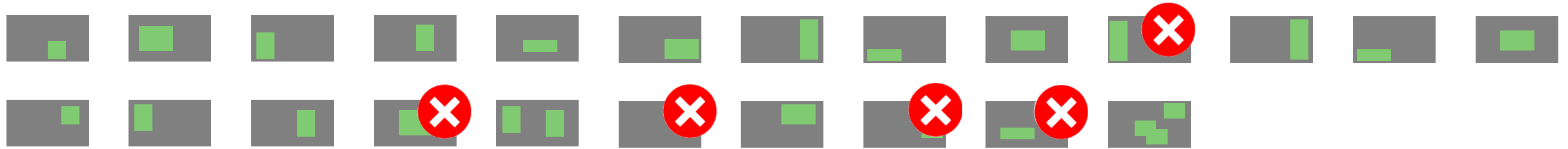
Schritt 1: Selektion betroffener Testfälle



Schritt 1: Selektion betroffener Testfälle



Schritt 2: Priorisierung selektierter Testfälle



Schritt 2: Priorisierung selektierter Testfälle



Pareto-Testing

Tests werden **unabhängig von Änderungen** ausgewählt.

95% der Fehler in 15% der Zeit ermittelt

Einmalige Messung der Coverage reicht aus (ggf. quartärlche Wiederholung)

Weniger Aufwand & breiter Anwendbar

Test-Impact-Analyse

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

90% der Fehler in 2% der Zeit ermittelt

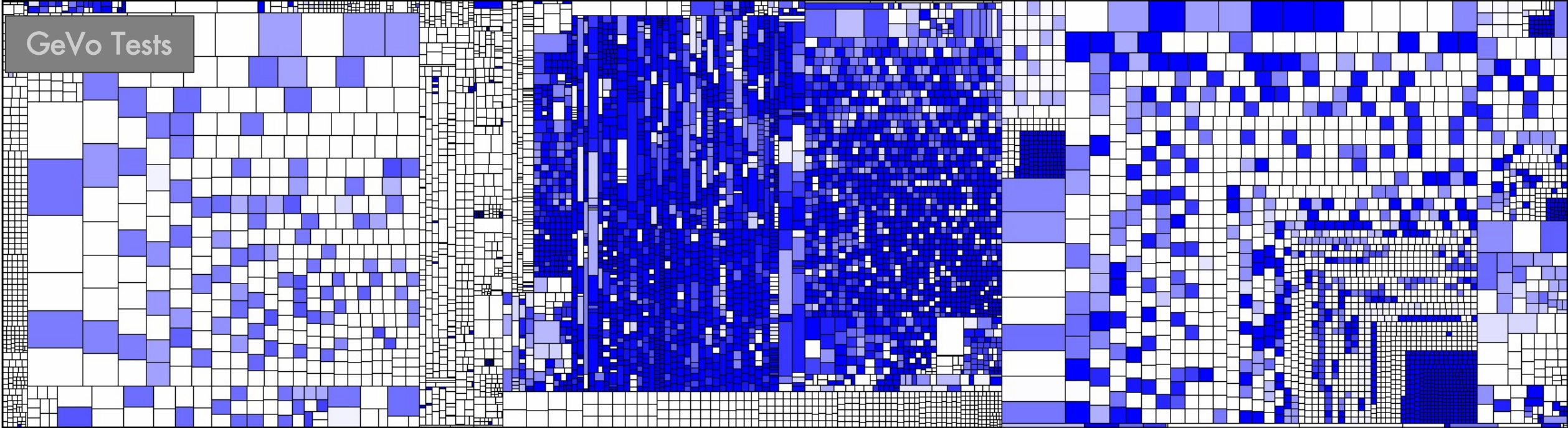
Erfordert kontinuierliche Messung der Coverage und Integration der Test-Auswahl in die CI / Testautomatisierung.

Stärkere Beschleunigung des Feedbacks (bei höherem Aufwand)

Was hilft der BVK?

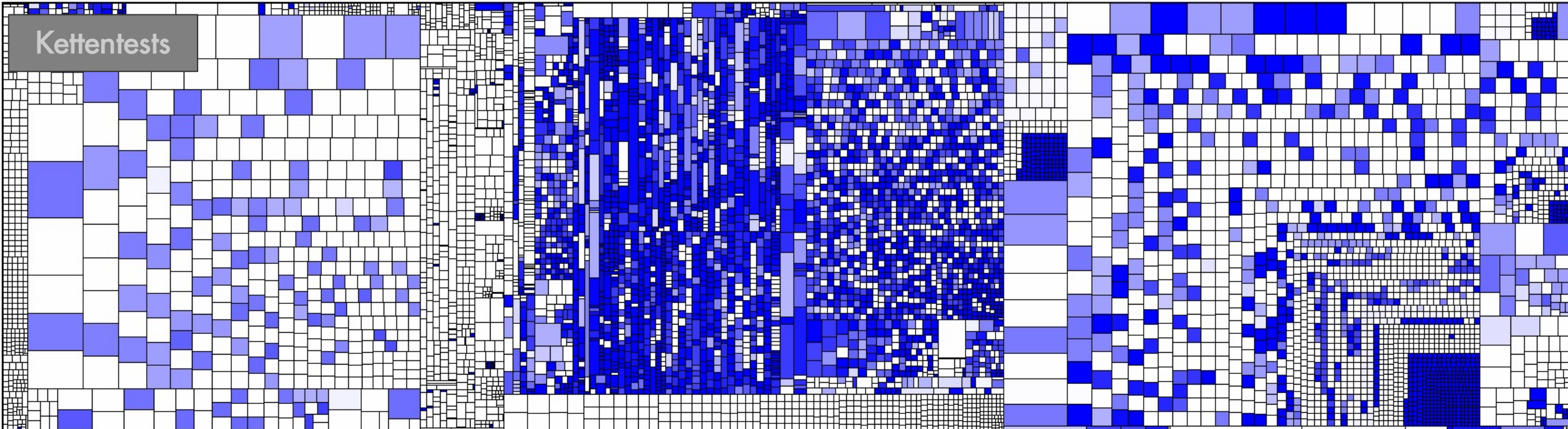


GeVo Tests



Es gibt noch weitere Testarten, die wir hier nicht vermessen haben, sowie generierten Code, der nicht getestet wird.

Kettentests



Es gibt noch weitere Testarten, die wir hier nicht vermessen haben, sowie generierten Code, der nicht getestet wird.

Historische Änderungen
Einmaliger Export aus CVS-Repository



Test-Impact-Analyse



Selektierte & priorisierte
Testfälle

Testfall-spezifische
Coverage



Wie effektiv sind die ausgewählten Tests?

Issue

VBO-6630

Erwartete Testfälle

GV_VBO_04_003F
 GV_VBO_04_004F
 GV_VBO_04_009F
 GV_VBO_04_309F
 TK-I-VersRiesterAnteil-001
 TK-I-VersRiesterAnteil-002
 TK-I-VersRiesterAnteil-003
 TK-I-VersRenteAendSchaedlVerw-001
 TK-I-VersRenteAendSchaedlVerw-002

VBO-6873

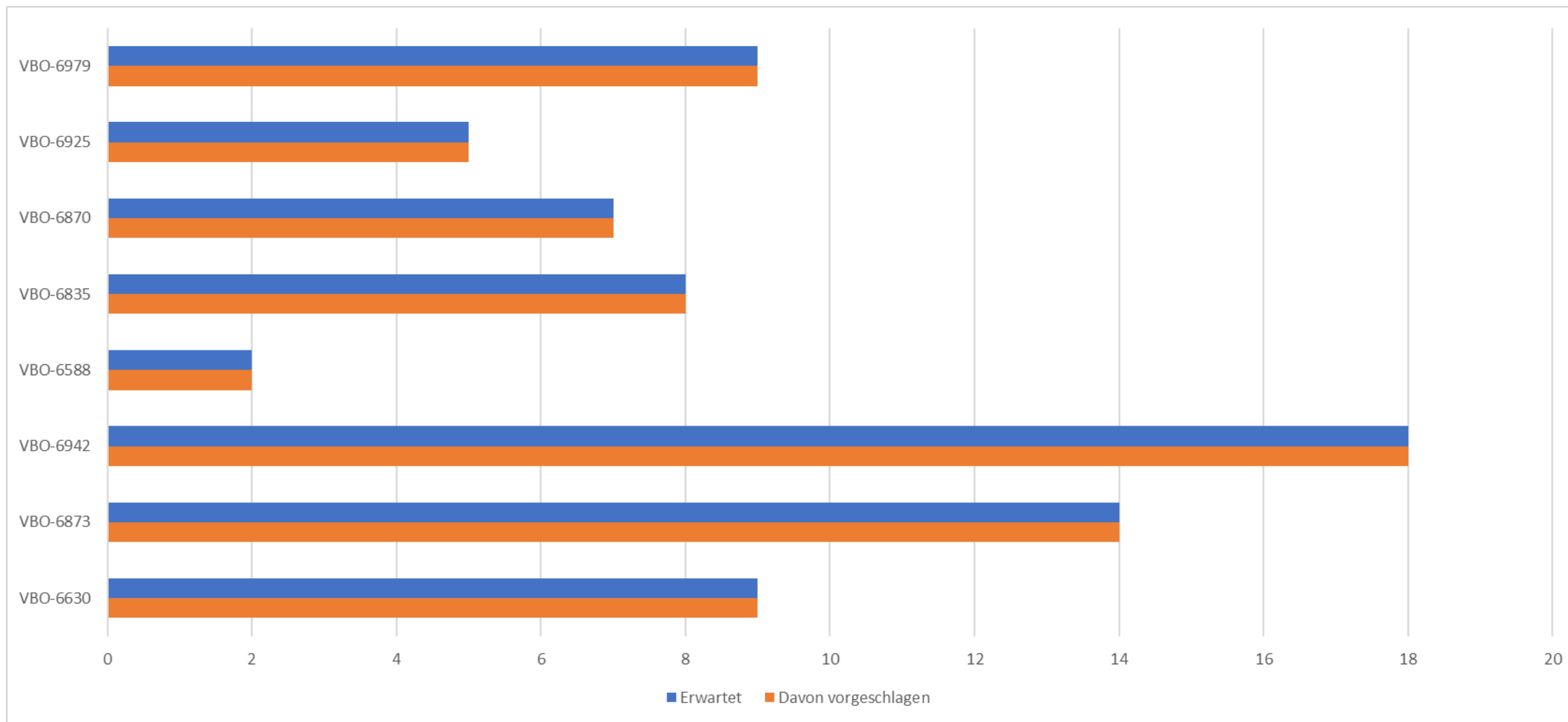
GV_VBO_05_041F
 GV_VBO_05_042F
 GV_VBO_05_342F
 TK-I-BESG-BU-TOD-001
 TK-I-BESG-BU-TOD-002
 TK-I-BESG-BU-TOD-003
 TK-I-BESG-BU-TOD-004
 TK-I-BESG-BU-ERW-AR-TOD
 TK-I-BESG-NZ-ANNU-004
 TK-I-VersAusgleichsrente-005
 TK-I-VersAusgleichsrente-006
 TK-I-VersAusgleichsrente-007
 TK-I-BESG-NM-WAISE-001
 TK-I-BESG-NM-WAISE-002

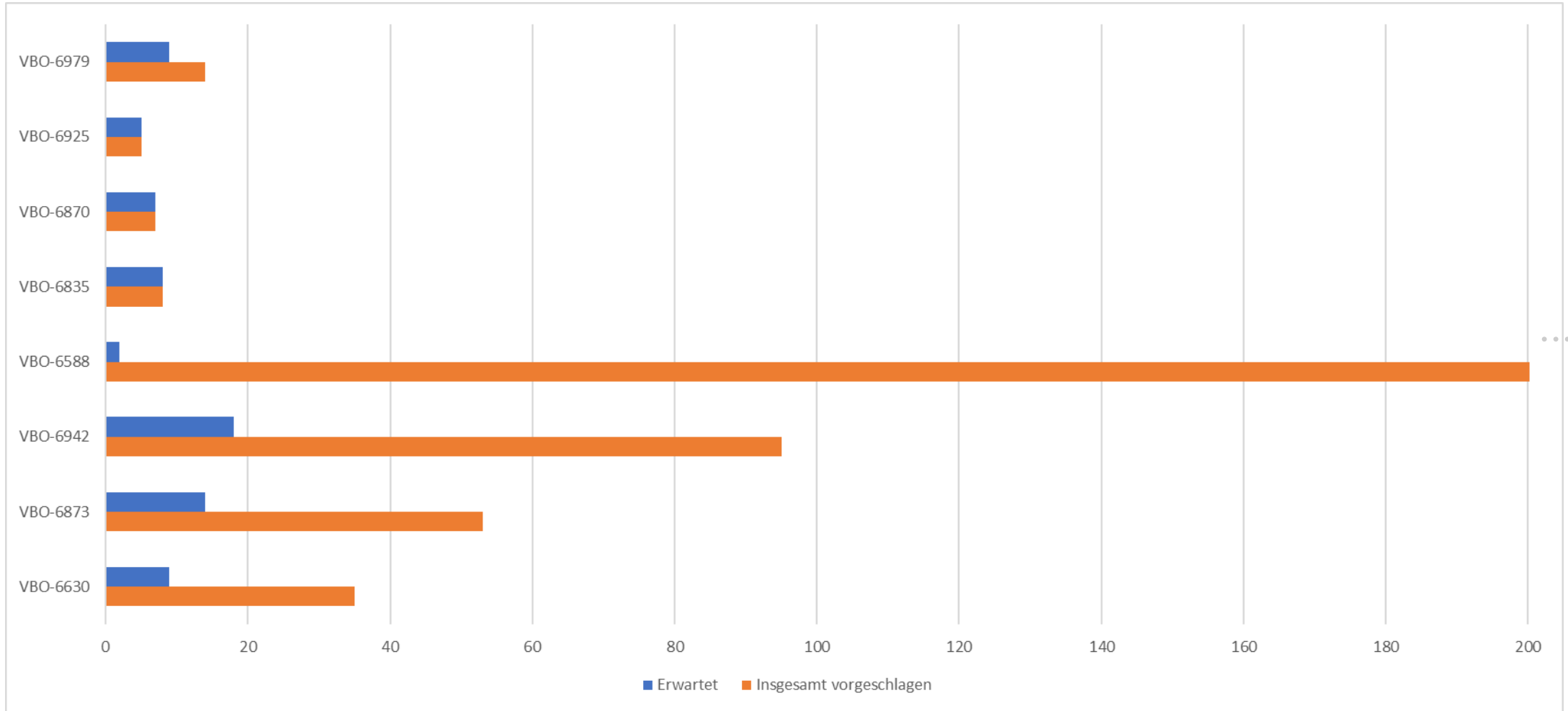
VBO-6942

GV_VBO_05_041F
 GV_VBO_05_042F
 GV_VBO_05_342F
 TK-I-BESG-NM-001
 TK-I-BESG-NM-002
 TK-I-BESG-NM-007
 TK-I-BESG-NM-WAISE-001
 TK-I-BESG-NM-WAISE-002
 TK-I-BESG-NM-WAISE-003
 TK-I-BESG-NM-WAISE-004
 TK-I-BESG-NM-WAISE-005
 TK-I-BESG-NM-WAISE-006
 TK-I-BESG-NM-WITWE-001
 TK-I-BESG-NM-WITWE-002
 TK-I-BESG-NM-WITWE-003
 TK-I-BESG-NM-WITWE-004
 TK-I-BESG-AR-TOD-001
 TK-I-BESG-AR-TOD-002

Impacted Tests

Test	Expected duration	Score	Reason
TS_GV_VBO_02_015F_Fehler_TF017	1s	5	Impacted
TS_GV_VBO_02_015F_Fehler_TF014	1s	3	Impacted
TS_GV_VBO_02_015F_Standard_TF029	2s	3	Impacted
TS_GV_VBO_02_015F_Standard_TF020	3s	2	Impacted
TS_GV_VBO_02_015F_Standard_TF012	3s	2	Impacted
TK_TK-I-NblstgBeitragserstattung-170	5s	1	Impacted
TK_TK-I-NblstgBeitragserstattung-162	6s	1	Impacted
TK_TK-I-NblstgBeitragserstattung-161	6s	1	Impacted
TK_TK-I-NblstgBeitragserstattung-160	37s	0	Impacted
TK_TK-I-Vers-004	44s	0	Impacted
TS_GV_VBO_02_015F_Standard_TF016	32s	0	Impacted
TS_GV_VBO_02_015F_Standard_TF011	32s	0	Impacted
TS_GV_VBO_02_015F_Standard_TF019	32s	0	Impacted
TS_GV_VBO_02_015F_Standard_TF015	32s	0	Impacted
TS_GV_VBO_02_015F_Standard_TF008	32s	0	Impacted
TS_GV_VBO_02_015F_Standard_TF006	32s	0	Impacted
TS_GV_VBO_02_015F_Standard_TF007	32s	0	Impacted
TS_GV_VBO_02_015F_Standard_TF010	32s	0	Impacted
TK_TK-I-Vers-013	57s	0	Impacted
TK_TK-I-Vers-008	1m 6s	0	Impacted
TK_TK-I-Vers-005	1m 52s	0	Impacted



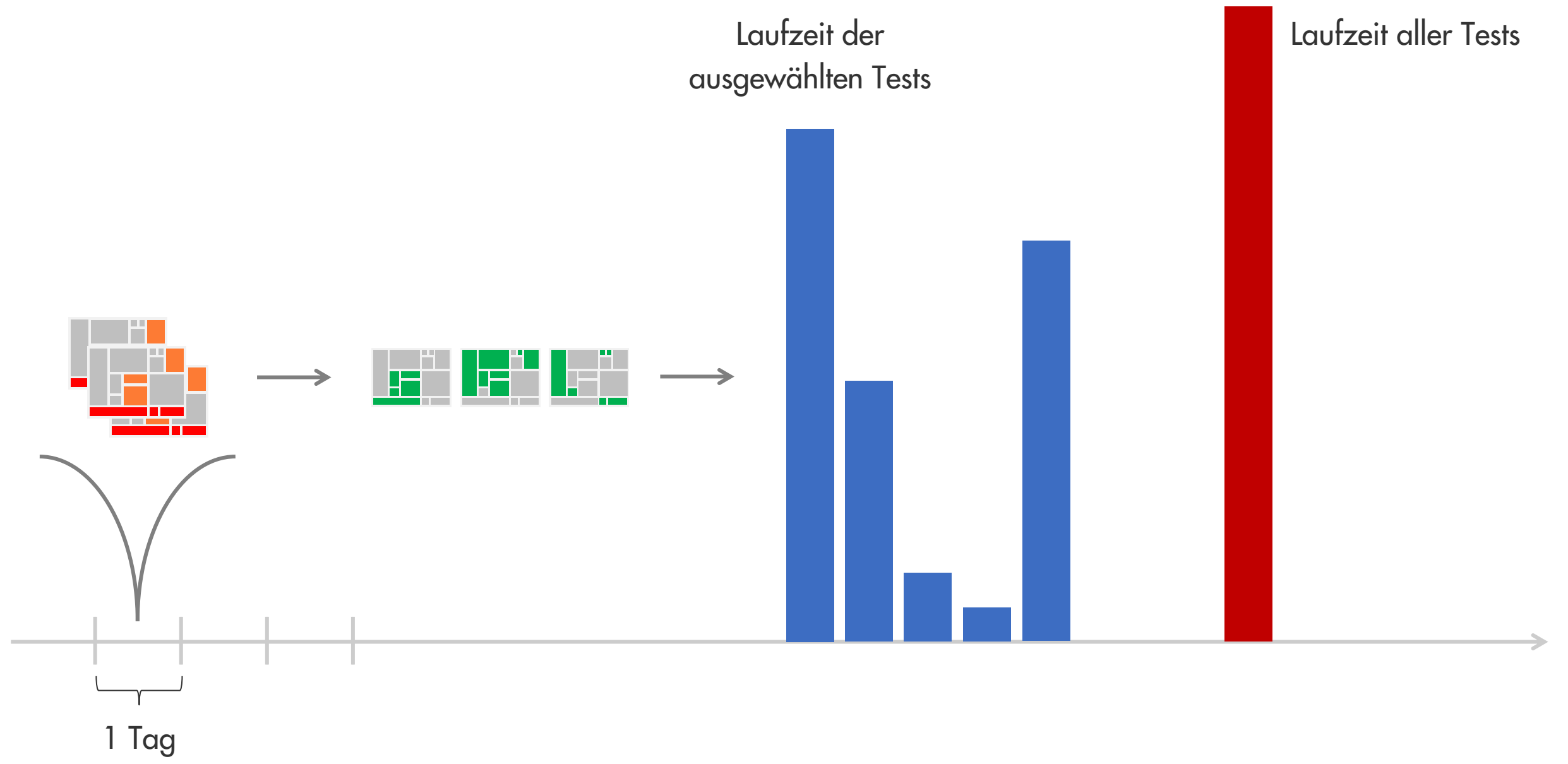


Teamscale schlägt an vielen Stellen noch weitere Tests vor, die von den Änderungen betroffen sind

Wie viel Testlaufzeit können wir sparen?

Laufzeit der
ausgewählten Tests

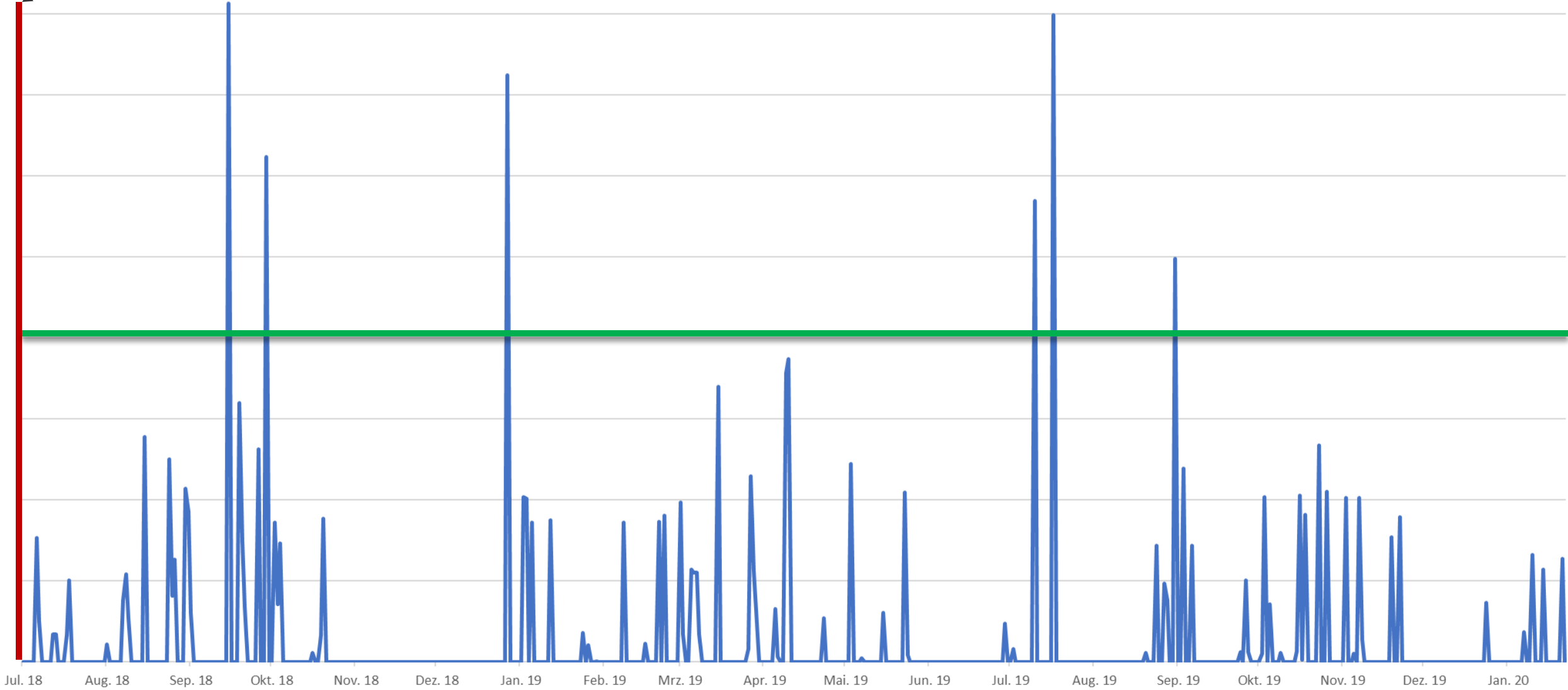
Laufzeit aller Tests



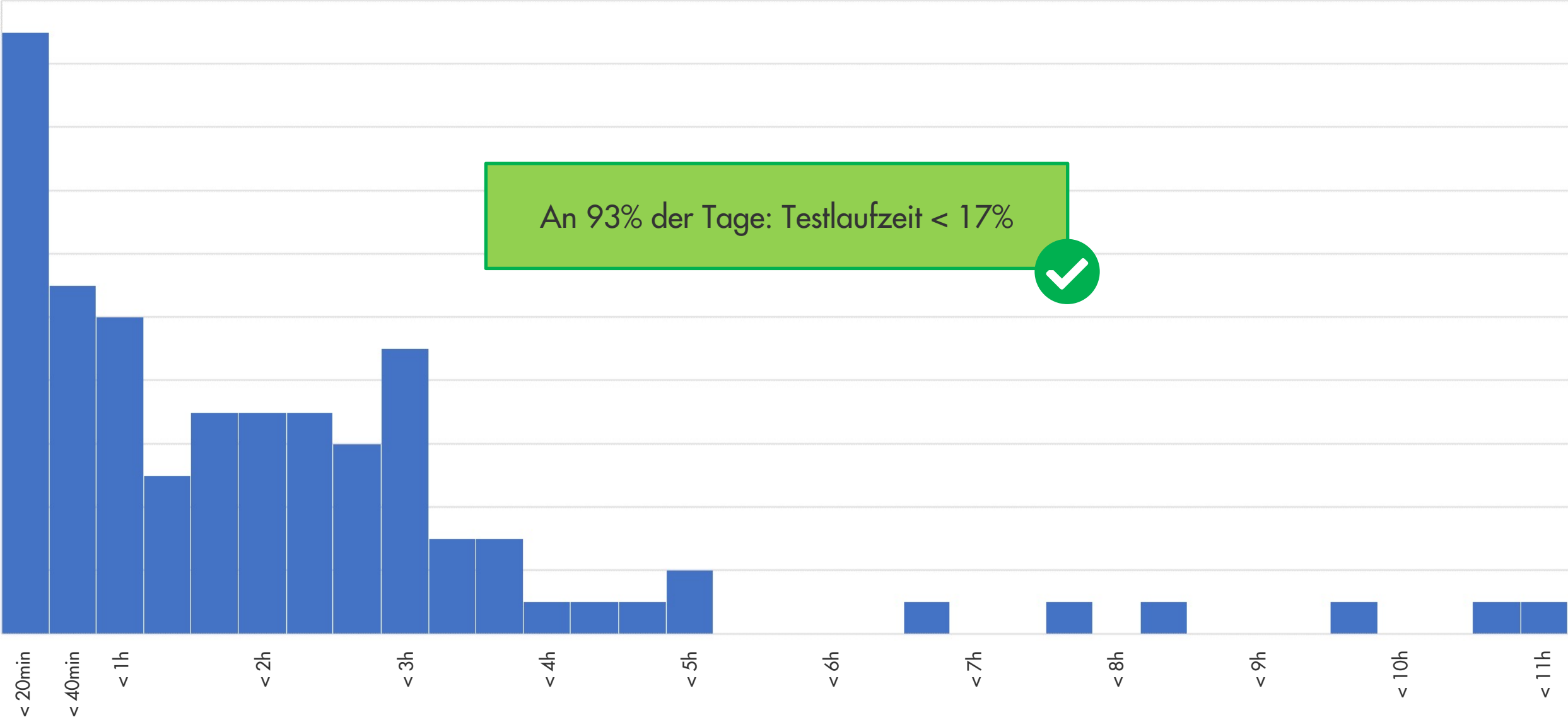
1 Tag

Impacted Test Laufzeit (tägliche Intervalle)

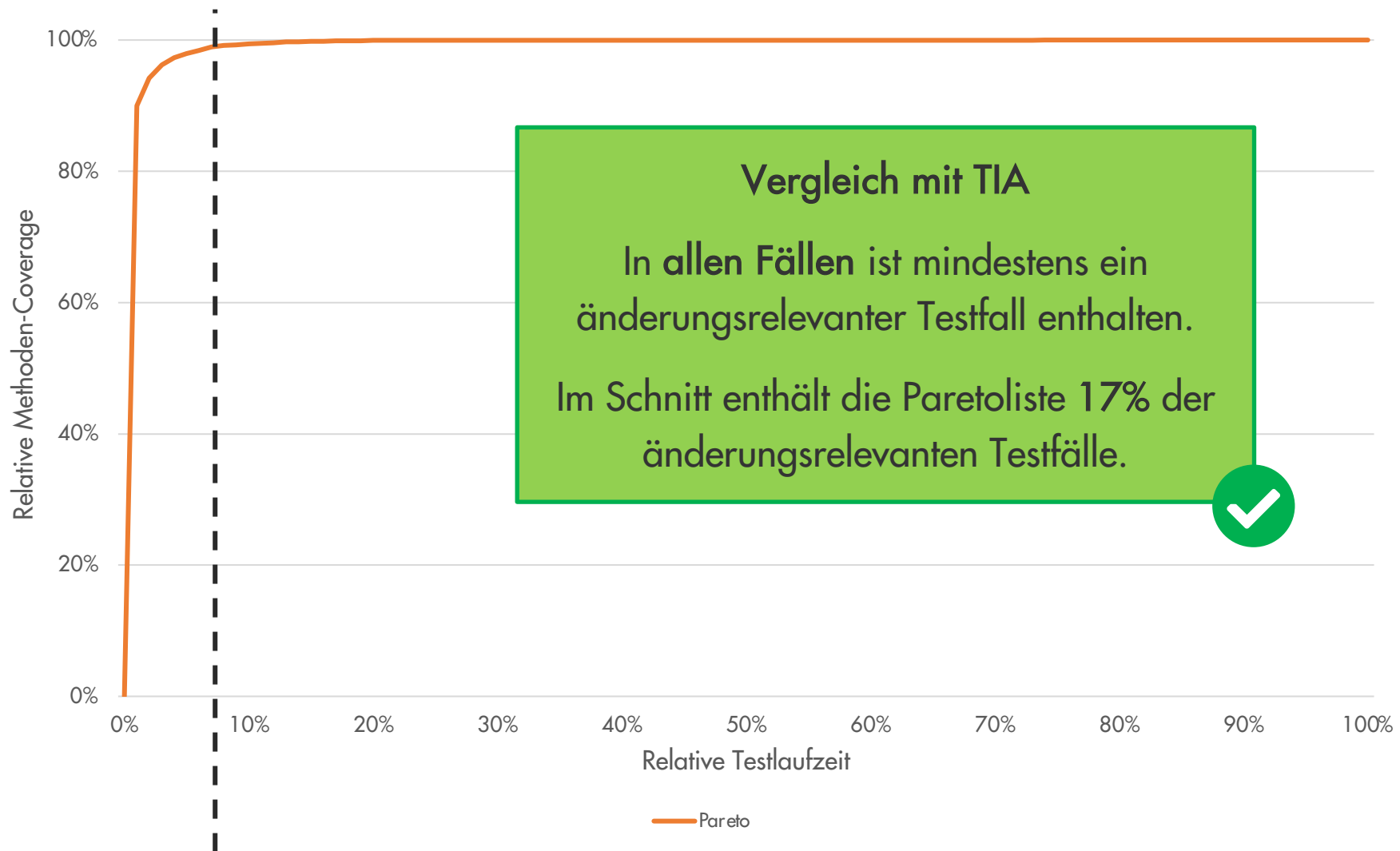
Alle Tests
ohne TIA



Laufzeitenhistogramm (tägliche Intervalle)



Nur Intervalle berücksichtigt, die Impacted Tests enthalten



Mit 1h Testlaufzeit erreichen wir bereits 99,2% relative Methoden-Coverage

Vorschlag für den Einsatz

Aktuell: nächtlich „**Re-run all**“

Feedback erst spät am nächsten Tag



Build

Test

Vorschlag für den Einsatz

Tagsüber **schnelles Feedback**: TIA Timeboxing

Erstes Feedback auf eingetragene Änderungen nach spätestens 2h



Build **Test max. 1h**

Vorschlag für den Einsatz

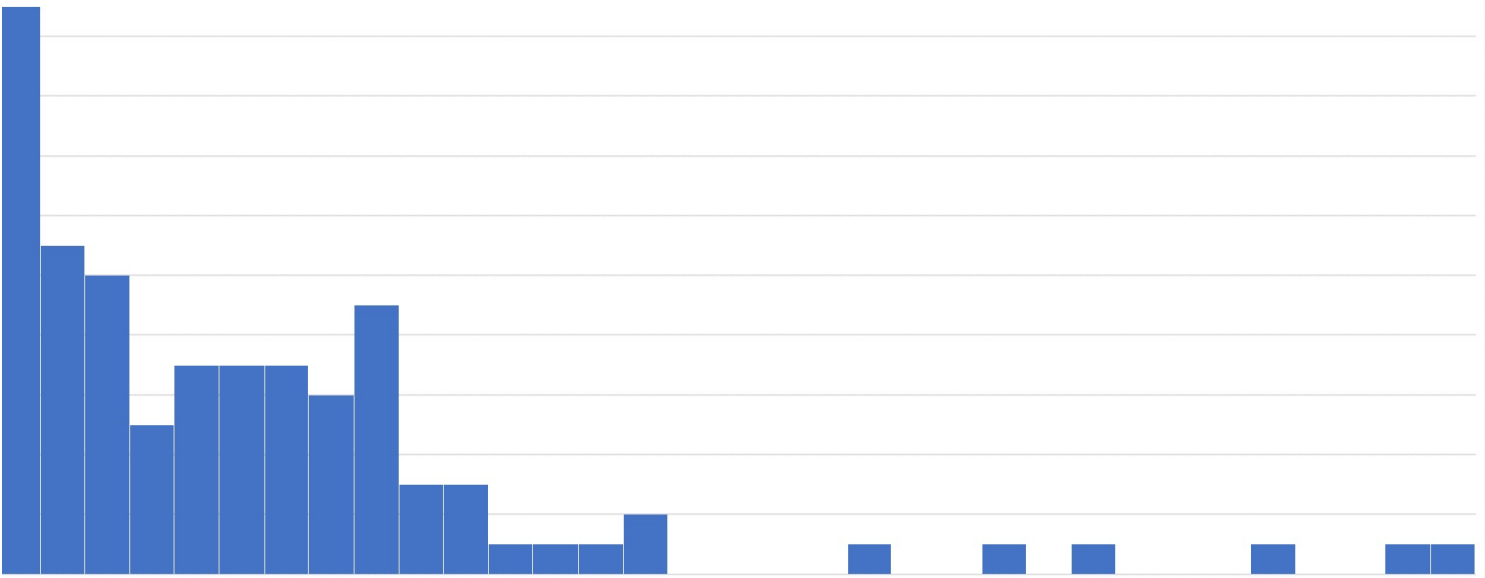
Tagsüber **schnelles Feedback**: TIA Timeboxing

Erstes Feedback auf eingetragene Änderungen nach spätestens 2h



Build **Test max. 1h**

Laufzeitenhistogramm (tägliche Intervalle)



Vorschlag für den Einsatz **nach der Git-Umstellung**

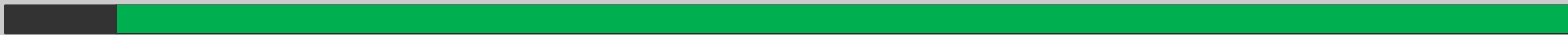
Tagsüber **schnelles Feedback**: TIA Timeboxing

Erstes Feedback auf eingetragene Änderungen nach spätestens 2h



Build **Test max. 1h**

Weiterhin: nächtlich „**Re-run all**“, um die testfallspezifische Coverage zu aktualisieren



Vorschlag für den Einsatz vor der Git-Umstellung: Pareto-Testliste

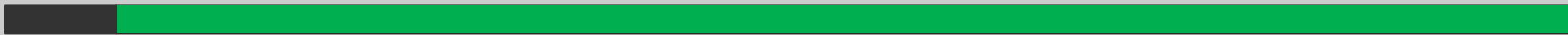
Tagsüber **schnelles Feedback**: Statisch berechnetes Testset

Erstes unvollständiges Feedback auf eingetragene Änderungen nach spätestens 2h



Build **Test 1h**

Weiterhin: nächtlich „**Re-run all**“, um die testfallspezifische Coverage zu aktualisieren



Monatlich: **Pareto-Liste aktualisieren**

Schnelles Feedback trotz langsamer Tests

Test-Impact-Analyse und Pareto-Optimierung
für historisch gewachsene Test-Suites

Dr. Elmar Jürgens & Fabian Streitl



Mittwoch, 1. Juni

10:30-12:00 Uhr

cqse.eu/workshop-8020-de-2206-gtd



Wednesday, June 29

17:00-18:30 CEST

cqse.eu/workshop-8020-en-2206-gtd



Vielen Dank! Ich freue mich auf Kommentare und Diskussion 🙌



Sven Amann · amann@cqse.eu · +49 172 1860063

CQSE GmbH
Centa-Hafenbrädl-Straße 59
81249 München

CQSE