



Test-Gap-Analyse

Wie erkennen wir ungetestete Änderungen, bevor sie ins Release gehen?
- Erfahrungen aus drei Jahren Praxiseinsatz

DWX Developer Week, 20. Juni 2016

Dr. Dennis Pagano

Über Mich

Forschung

- Analyse von Open Source Communities
- Code, Test- und Kommunikationsartefakten
- Kontinuierliches Einbeziehen von Nutzern in die Entwicklung



Beratung

- Qualitätsbewertung und Qualitätscontrolling
- Software und Systems Engineering



Entwicklung

- Teamscale: Test-Gap-Analyse





Agenda

- **Motivation**
- Test-Gap-Analyse
- Demo
- Erfahrungen aus der Praxis
- Ausblick



Optimierung Kosten & Nutzen von Testaktivitäten

Fehlerfolgekosten

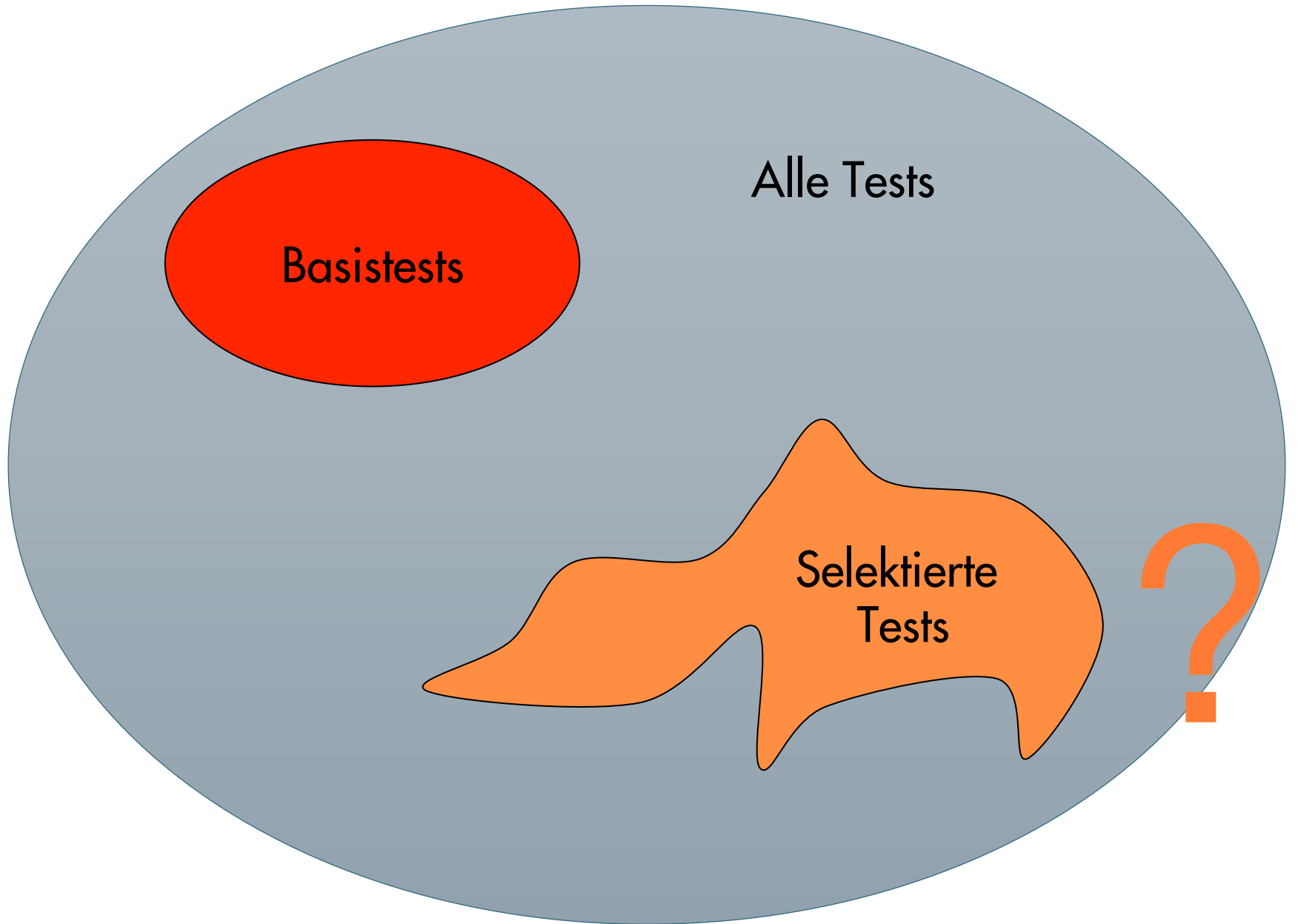
×

Fehlerwahrscheinlichkeit

>

Testkosten





Alle Tests

Basistests

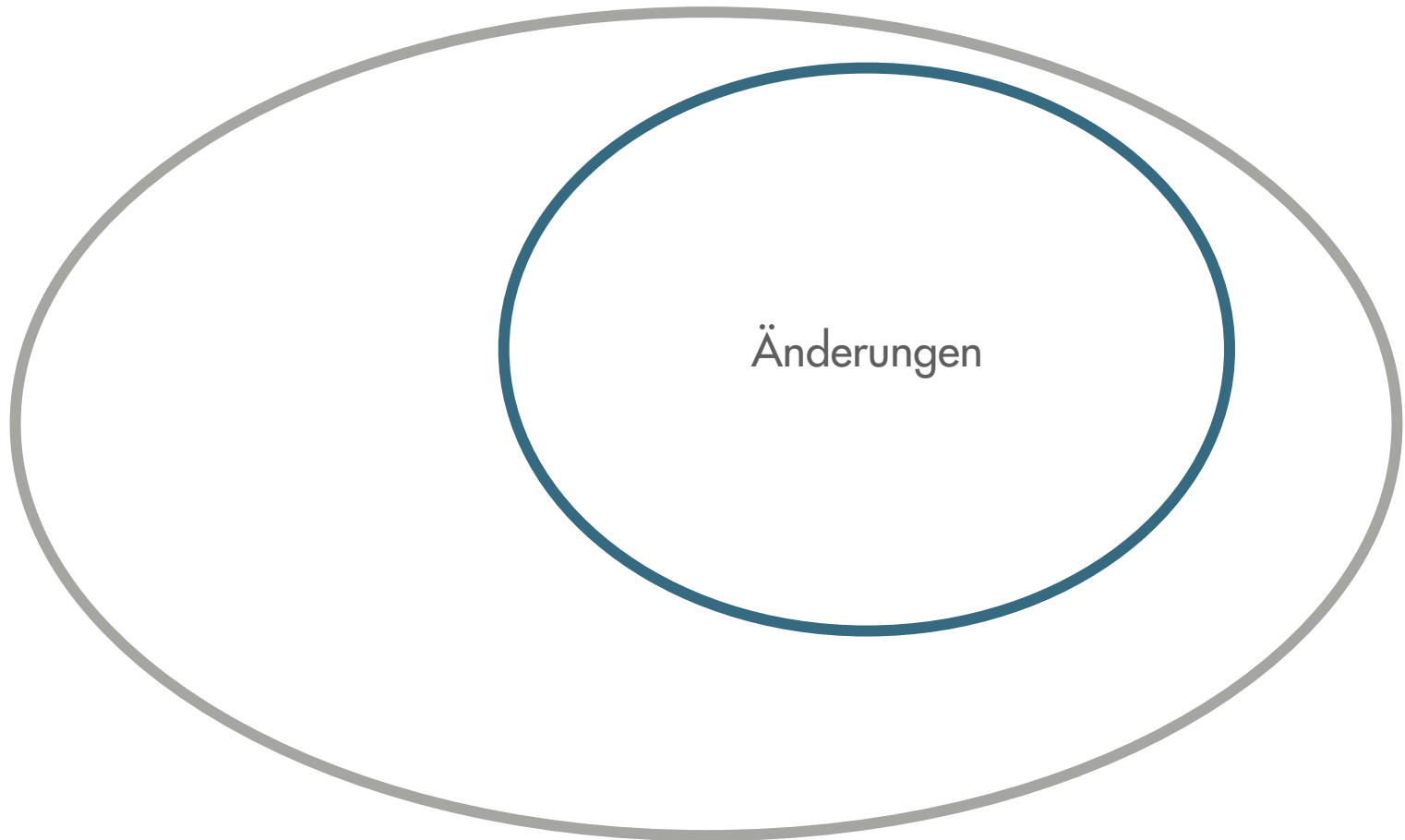
Selektierte
Tests



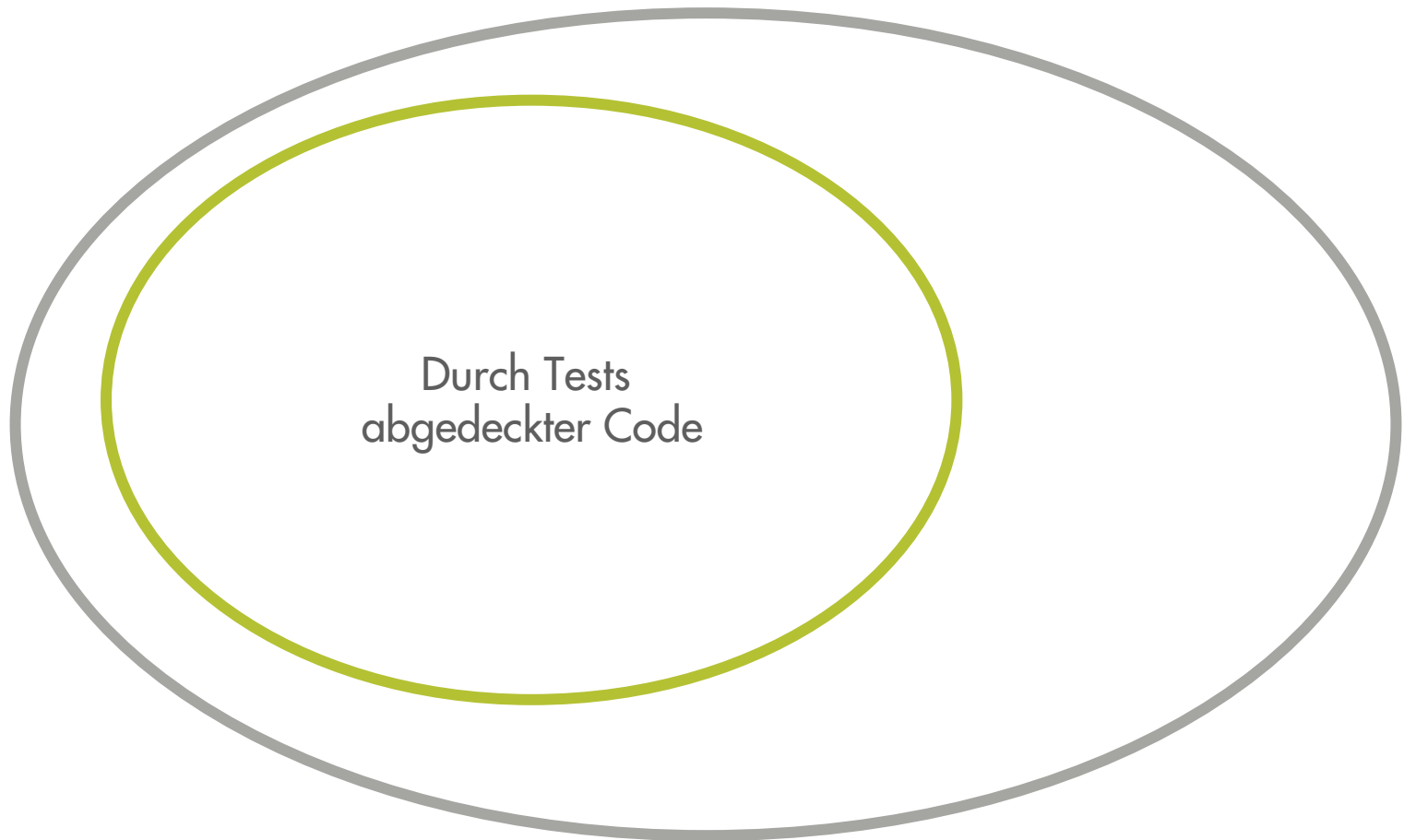
Worauf Testaktivitäten fokussieren?

**→ Wo finden wir
die meisten Bugs?**

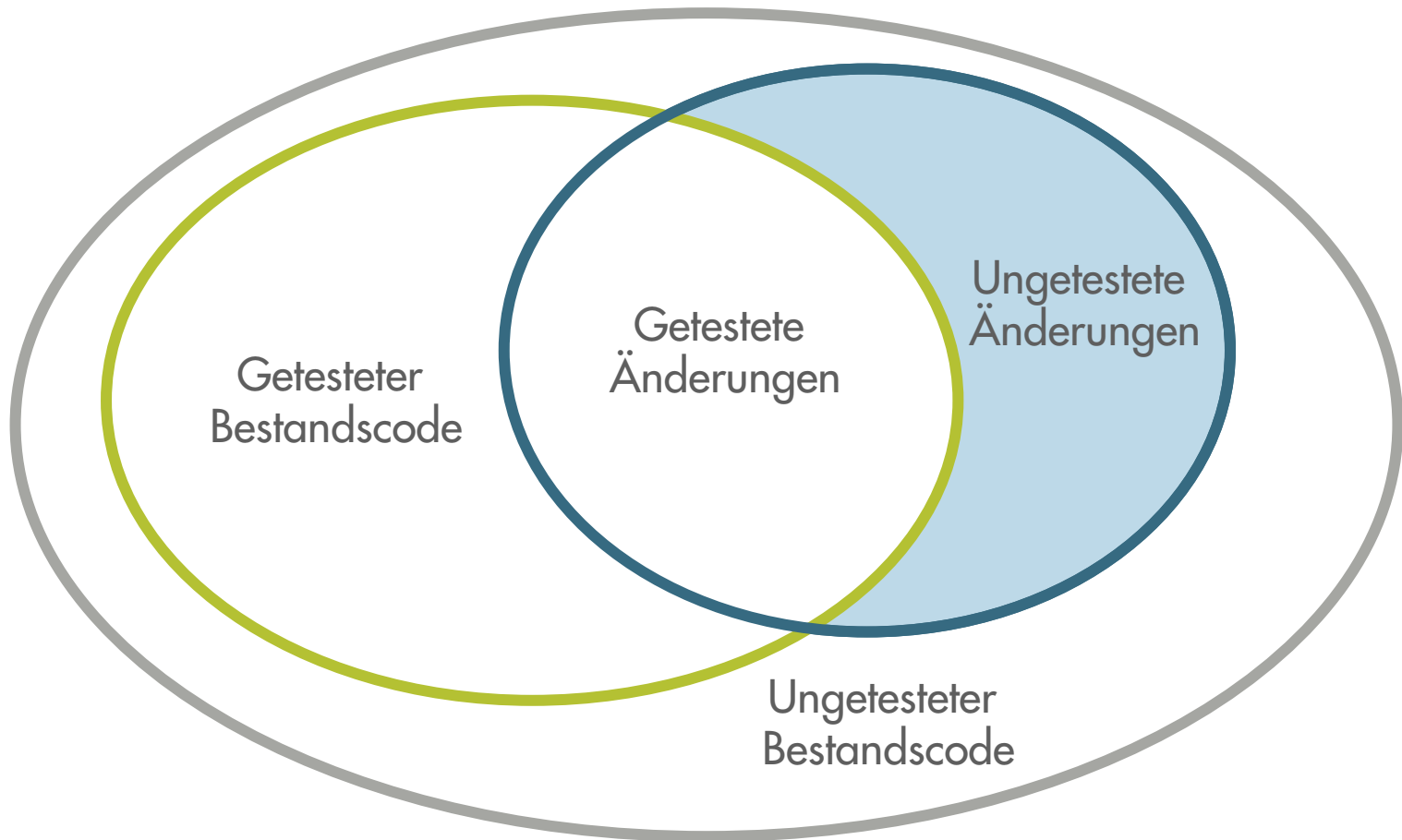
Hintergrund



Hintergrund



Hintergrund



Hintergrund

Studie:
Woher kommen Feldfehler?
C# System @ Munich Re

**Fünfmal so hohe
Fehlerwahrscheinlichkeit
in ungetesteten Änderungen**

**Die Hälfte der Änderungen
wurde nicht getestet**

Ungetestete
Änderungen

getesteter
Handcode

Änderungen sind guter Fehlerprediktor

Nagappan, 2005

- Code Churn guter Prädiktor für Feldfehler
- Evaluiert auf *Windows Server 2003 (45 MLOC)*

Hassan, Holt 2005

- Most frequently/recently changed guter Fehlerprädiktor
- Eval: *NetBSD, FreeBSD, OpenBSD, Postgres, KDE, Koffice*

Eder, Hauptmann, Junker, Juergens, Vass, Prommer, 2013

- Feldfehlerwahrscheinlichkeit 5x so hoch in geändertem Code
- Evaluiert auf *betr. Informationssystem @ Munich Re*

Bugzilla - Bug List

Home | New | Browse | Search | Search [?] | Reports | Preferences | Administration | Help | Log out juergens@in.tum.de

Wed Aug 28 2013 11:45:48 CEST
[quip?](#)

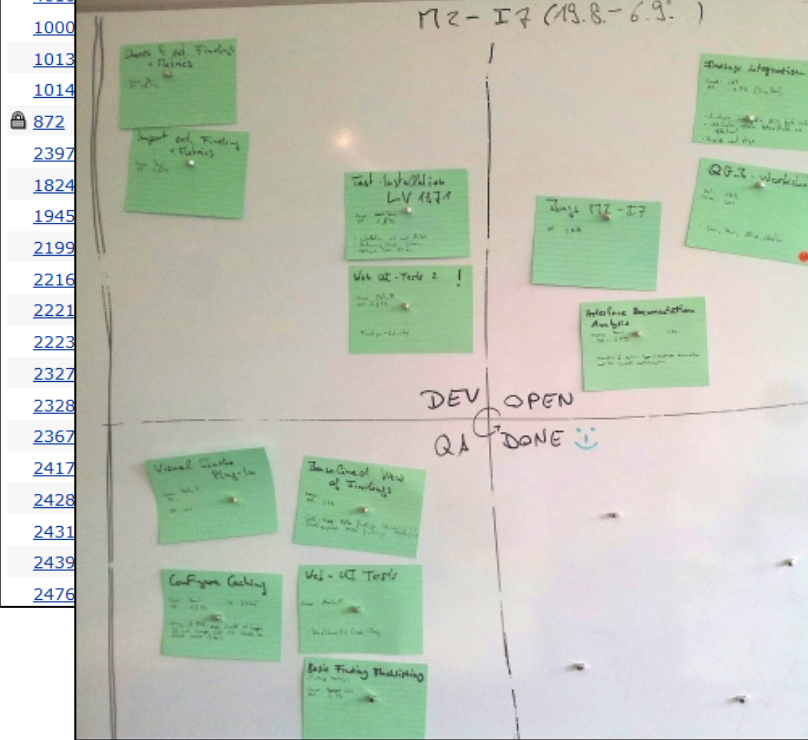
[Hide Search Description](#)

Status: UNCONFIRMED, NEW, ASSIGNED, REOPENED Product: clone Component: clone

83 bugs found.

ID ▲ Summary

- [3382](#) Remove old filtering framework
- [4608](#) Improve progress monitoring for clone detection
- [3613](#) Make MetricOverview more flexible w.r.t. to database table names
- [3811](#) Incorrect offset reported by two-phased scanners
- [4610](#) Clone compare view does not open clone pair at correct position



Übersicht
Aktivität
Roadmap
Tickets
Neues Ticket
Projektarchiv
Konfiguration

Tickets

Filter

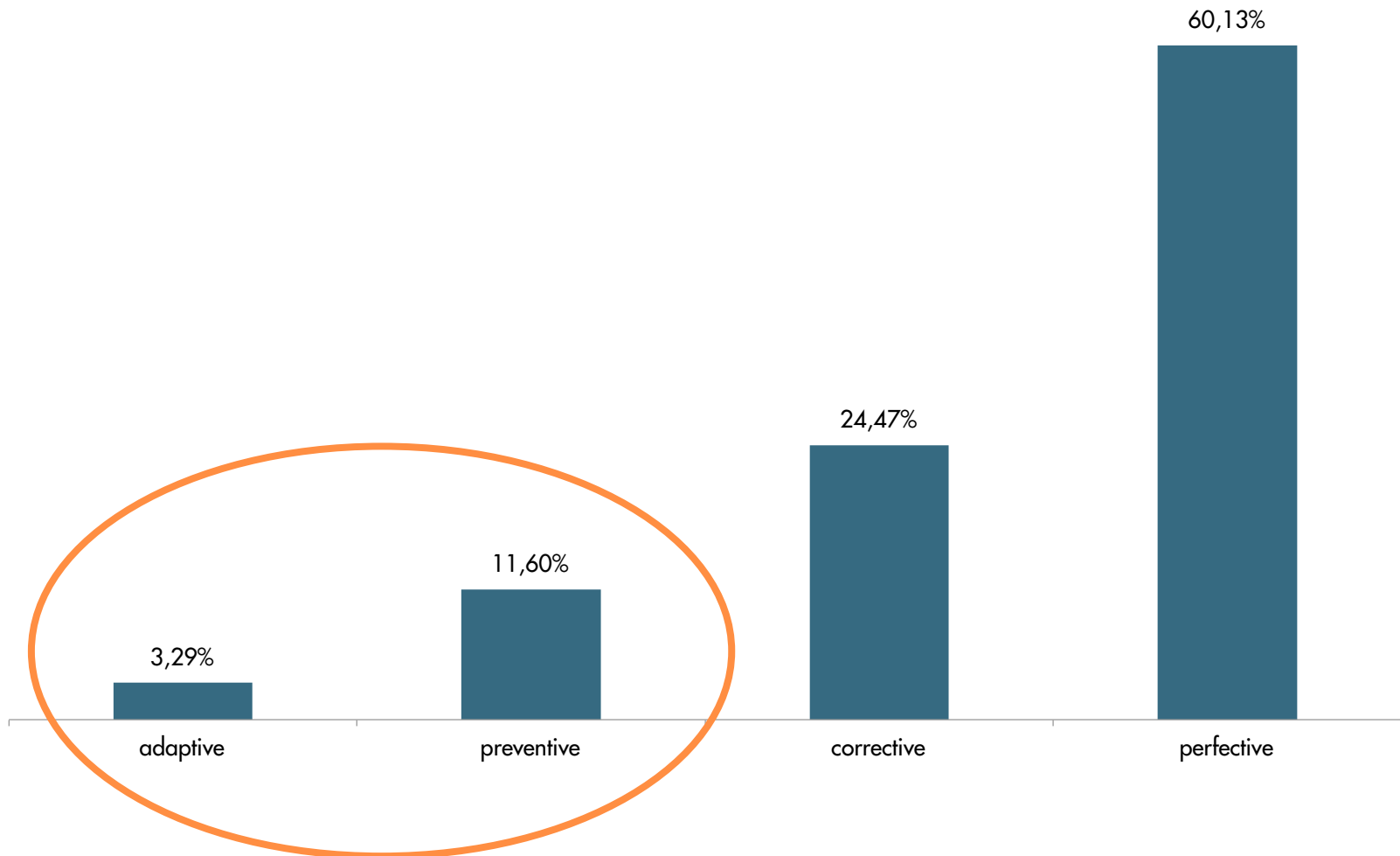
Status offen ▾

Optionen

Anwenden Zurücksetzen Speichern

#	Tracker	Status	Priorität	Them
<input type="checkbox"/> 5489	corrective	New	Normal	Cannot install ConQAT from update site on Kepler
<input type="checkbox"/> 5488	adaptive	New	Normal	Should be able to view "own" findings
<input type="checkbox"/> 5487	uncategorized	Assigned	Normal	Upload/Change of user image broken
<input type="checkbox"/> 5486	uncategorized	New	Normal	Allow for branded dists
<input type="checkbox"/> 5485	perfective	New	Normal	Enable hiding of individual findings in the IDE (e.g. after they are resolved)
<input type="checkbox"/> 5484	corrective	New	High	Visual Studio crashes after editing on findings
<input type="checkbox"/> 5483	uncategorized	Yellow	Normal	Cleanup com.teamscale.ide.prefs
<input type="checkbox"/> 5482	uncategorized	New	Normal	Filter structural clones
<input type="checkbox"/> 5481	adaptive	New	Normal	Integrate Sharepoint Analysis in MR Dashboard
<input type="checkbox"/> 5480	uncategorized	New	Normal	Teamscale connection preference page not working in Eclipse 4.3 on Mac
<input type="checkbox"/> 5479	preventive	Red	Normal	Benchmarking and selection of storage systems
<input type="checkbox"/> 5478	uncategorized	Yellow	Normal	Usability: "Open in Web Browser" should use the browser from preferences
<input type="checkbox"/> 5477	corrective	New	Normal	Fix refresh problem when setting a dashboard as default
<input type="checkbox"/> 5475	perfective	Assigned	Normal	Create UI test for "show history" in code perspective
<input type="checkbox"/> 5474	adaptive	Assigned	Normal	Enable mapping of baseline to head resources when packages are split
<input type="checkbox"/> 5473	perfective	Red	Urgent	Connection problems using svn+ssh
<input type="checkbox"/> 5472	perfective	Accepted	Normal	Display baselines in charts (as small vertical lines)
<input type="checkbox"/> 5471	perfective	Accepted	Normal	Support selection of explicit baseline for filtering
<input type="checkbox"/> 5470	perfective	Accepted	Normal	Service documentation shall also provide more detailed documentation on q
<input type="checkbox"/> 5469	perfective	Accepted	Normal	Adding/removing metrics to chart widget changes colors of other metrics
<input type="checkbox"/> 5468	perfective	Accepted	Normal	Simple metric widget shall support explicit baseline for delta value calculation
<input type="checkbox"/> 5467	corrective	Accepted	High	The delta not found in the store!
<input type="checkbox"/> 5466	corrective	Accepted	Normal	Problems when tracking TODOs which have changed line breaks

CRs in ConQAT



Auf Basis von ca. 2000 CRs für ConQAT

Bugzilla - Bug List

[Home](#) | [New](#) | [Browse](#) | [Search](#) | [?] | [Reports](#) | [Preferences](#) | [Administration](#) | [Help](#) | [Log out](#) juergens@in.tum.de

Wed Aug 28 2013 11:45:48 CEST

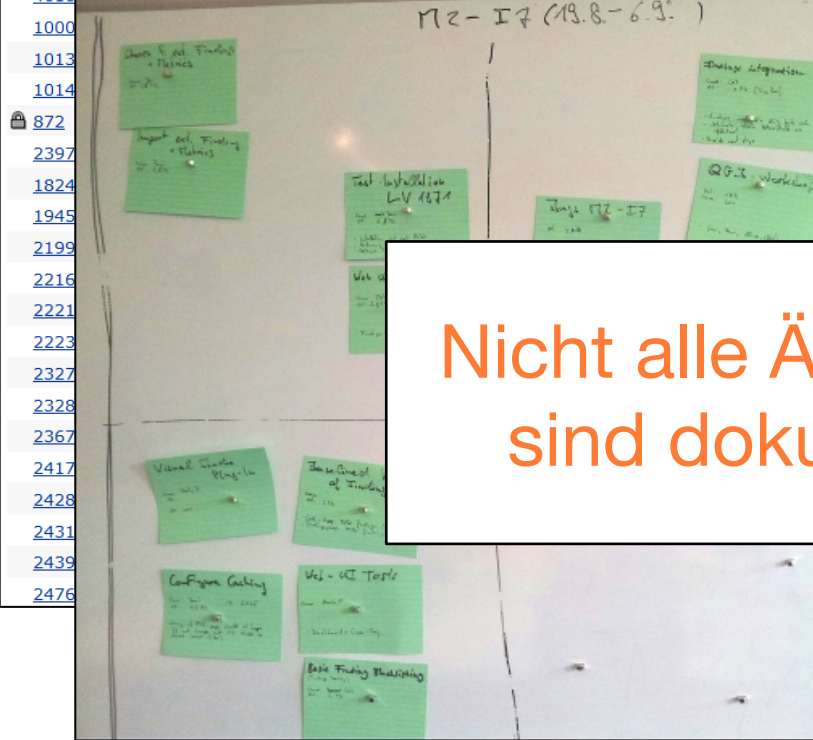
[quip?](#)

[Hide Search Description](#)

Status: UNCONFIRMED, NEW, ASSIGNED, REOPENED Product: clone Component: clone

83 bugs found.

- [ID](#) ▲ [Summary](#)
- [3382](#) Remove old filtering framework
- [4608](#) Improve progress monitoring for clone detection
- [3613](#) Make MetricOverview more flexible w.r.t. to database table names
- [3811](#) Incorrect offset reported by two-phased scanners
- [4610](#) Clone compare view does not open clone pair at correct position



Nicht alle Änderungen sind dokumentiert!

ConQAT Tickets

Übersicht Aktivität Roadmap Tickets Neues Ticket Projektarchiv Konfiguration

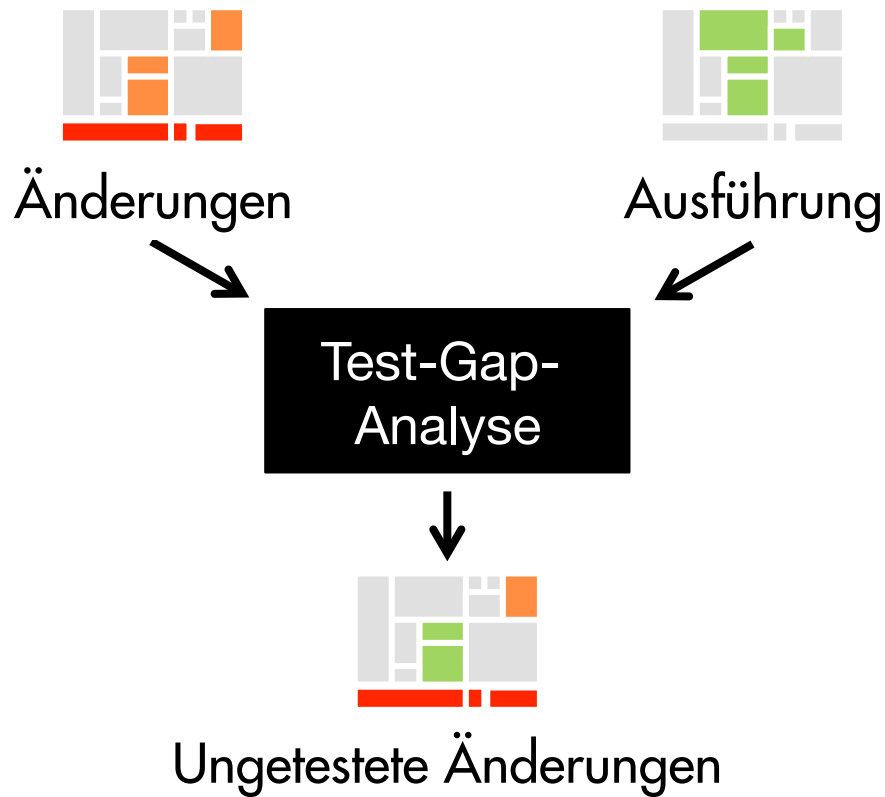
Filter: Status: Status: Optionen:

#	Tracker	Status	Priorität	Them
5489	corrective	New	Normal	Cannot install ConQAT from update site on Kepler
5488	adaptive	New	Normal	Should be able to view "own" findings
5487	uncategorized	Assigned	Normal	Upload/Change of user image broken
5486	uncategorized	New	Normal	Allow for branded dists
5485	perfective	Accepted	Normal	Enable hiding of individual findings in the IDE (e.g. after they are resolved)
5484	perfective	Accepted	High	Visual Studio crashes after editing on findings
5483	perfective	Accepted	Normal	Cleanup com.teamscale.ide.prefs
5482	perfective	Accepted	Normal	Filter structural clones
5481	perfective	Accepted	Normal	Integrate Sharepoint Analysis in MR Dashboard
5480	perfective	Accepted	Normal	Teamscale connection preference page not working in Eclipse 4.3 on Mac
5479	perfective	Accepted	Normal	Benchmarking and selection of storage systems
5478	perfective	Accepted	Normal	Usability: "Open in Web Browser" should use the browser from preferences
5477	perfective	Accepted	Normal	Fix refresh problem when setting a dashboard as default
5476	perfective	Accepted	Normal	Create UI test for "show history" in code perspective
5475	perfective	Accepted	Normal	Enable mapping of baseline to head resources when packages are split
5474	adaptive	Assigned	Normal	Connection problems using svn+ssh
5473	perfective	Red	Urgent	Connection problems using svn+ssh
5472	perfective	Accepted	Normal	Display baselines in charts (as small vertical lines)
5471	perfective	Accepted	Normal	Support selection of explicit baseline for filtering
5470	perfective	Accepted	Normal	Service documentation shall also provide more detailed documentation on q
5469	perfective	Accepted	Normal	Adding/removing metrics to chart widget changes colors of other metrics
5468	perfective	Accepted	Normal	Simple metric widget shall support explicit baseline for delta value calculation
5467	corrective	Accepted	High	The delta not found in the store!
5466	corrective	Accepted	Normal	Problems when tracking TODOs which have changed line breaks

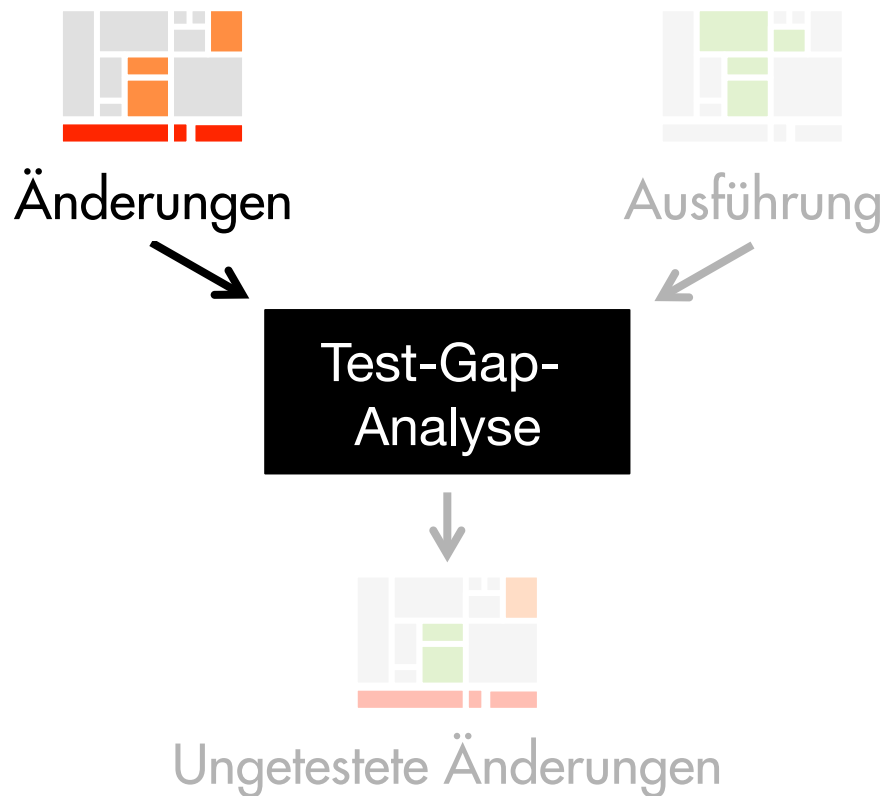
Agenda

- Motivation
- **Test-Gap-Analyse**
- Demo
- Erfahrungen aus der Praxis
- Ausblick

Die Test-Gap-Analyse



Änderungserkennung



- Gegenüber Referenzstand
- Akkumulierte Analyse
- Nur funktionale Änderungen
 - Keine Kommentare
 - Keine Umbenennungen
 - Keine Verschiebungen
 - Keine Refactorings

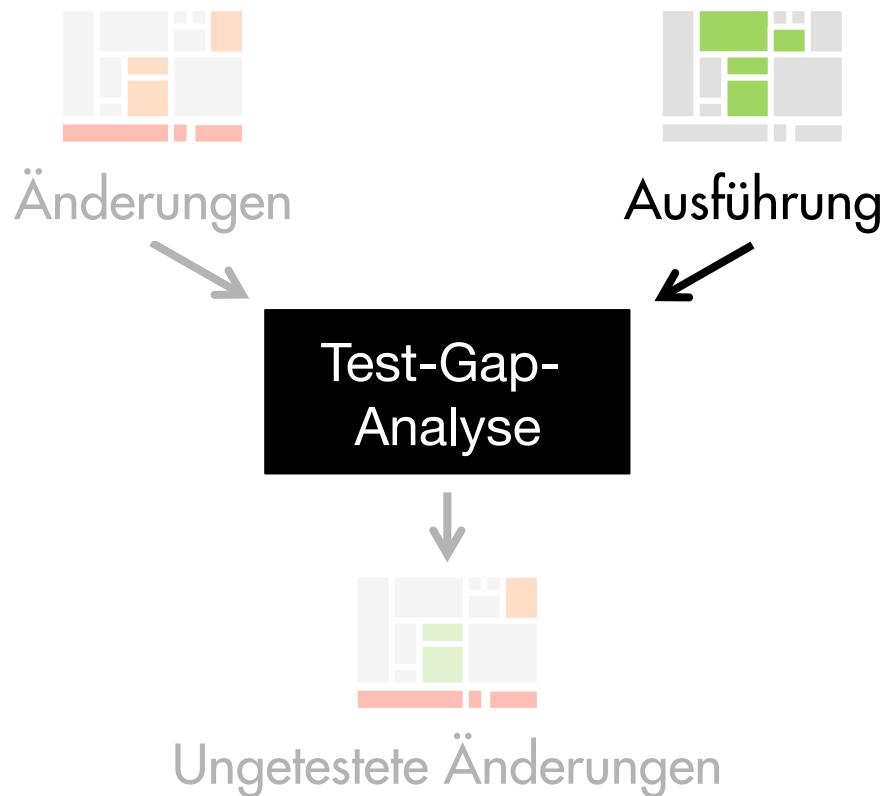
NamespaceRenames.java 8737

```
23
24
25
26 /* package */class NamespaceRenames {
27
28     /** Maps from old name fragment to new name fragment */
29     public Set<ImmutablePair<String, String>> namespaceRenames = new HashSet<
30
31     /**
32      * Computes a rename rule from an old and a new namespace name based on a
33      * type name correspondence. <code>
34      * For name1 = a.oldnamespace.b.c.D and name2 = a.newnamespace.b.c.D, res
35      * </code>
36      * */
37     public static ImmutablePair<String, String> computeRenameRule(
38         String fqTypeName1, String fqTypeName2, String separator) {
39
40         if (fqTypeName1.equals(fqTypeName2)) {
41             return null;
42         }
43
44         String commonSuffix = StringUtils.longestCommonSuffix(fqTypeName1,
45             fqTypeName2);
46         if (StringUtils.isEmpty(commonSuffix)) {
47             return null;
48         }
49
50         int separatorPosition = commonSuffix.indexOf(separator);
51         if (separatorPosition != -1) {
52             commonSuffix = commonSuffix.substring(separatorPosition);
53         }
54
55         String from = StringUtils.stripSuffix(commonSuffix, fqTypeName1);
56         String to = StringUtils.stripSuffix(commonSuffix, fqTypeName2);
57
58         return new ImmutablePair<String, String>(from, to);
59     }
60
61
62
63
64
```

NamespaceRenames.java 8727

```
24
25
26 /* package */class NamespaceRenames {
27
28     /** Maps from old name fragment to new name fragment */
29     public Set<ImmutablePair<String, String>> namespaceRenames = new Has
30
31     // TODO (LH) Doc vs. method name : 'Compute' vs. 'find'
32     /**
33      * Computes a rename rule from an old and a new namespace name based
34      * type name correspondence. <code>
35      * For name1 = a.oldnamespace.b.c.D and name2 = a.newnamespace.b.c.D
36      * </code>
37      * */
38     // TODO (LH) Please reflect in identifier names that they refer to
39     // names
40     public static ImmutablePair<String, String> findRenameRule(String na
41         String name2, String separator) {
42
43         if (name1.equals(name2)) {
44             return null;
45         }
46
47         // TODO (LH) Looks like StringUtils is missing a 'longestCommonS
48         String commonSuffix = reverse(StringUtils.longestCommonPrefix(
49             reverse(name1), reverse(name2)));
50         if (StringUtils.isEmpty(commonSuffix)) {
51             return null;
52         }
53
54         int separatorPosition = commonSuffix.indexOf(separator);
55         if (separatorPosition != -1) {
56             commonSuffix = commonSuffix.substring(separatorPosition);
57         }
58
59         // TODO (LH) Please reflect in identifier names that these are
60         // namespaces
61         String from = StringUtils.stripSuffix(commonSuffix, name1);
62         String to = StringUtils.stripSuffix(commonSuffix, name2);
63
64         return new ImmutablePair<String, String>(from, to);
65     }
66
```

Auswertung der Ausführung



- Granularität: Methoden
- Manuelle oder Unit-Tests
- Auf Server oder Clients
- Minimaler Overhead
- Plattformspezifisch

conQAT

dotCover 2

 Clover

NCOVER
Measurable Quality

 EMMA

 Microsoft
Visual Studio
Performance Tools

 JACOCO
Java Code Coverage

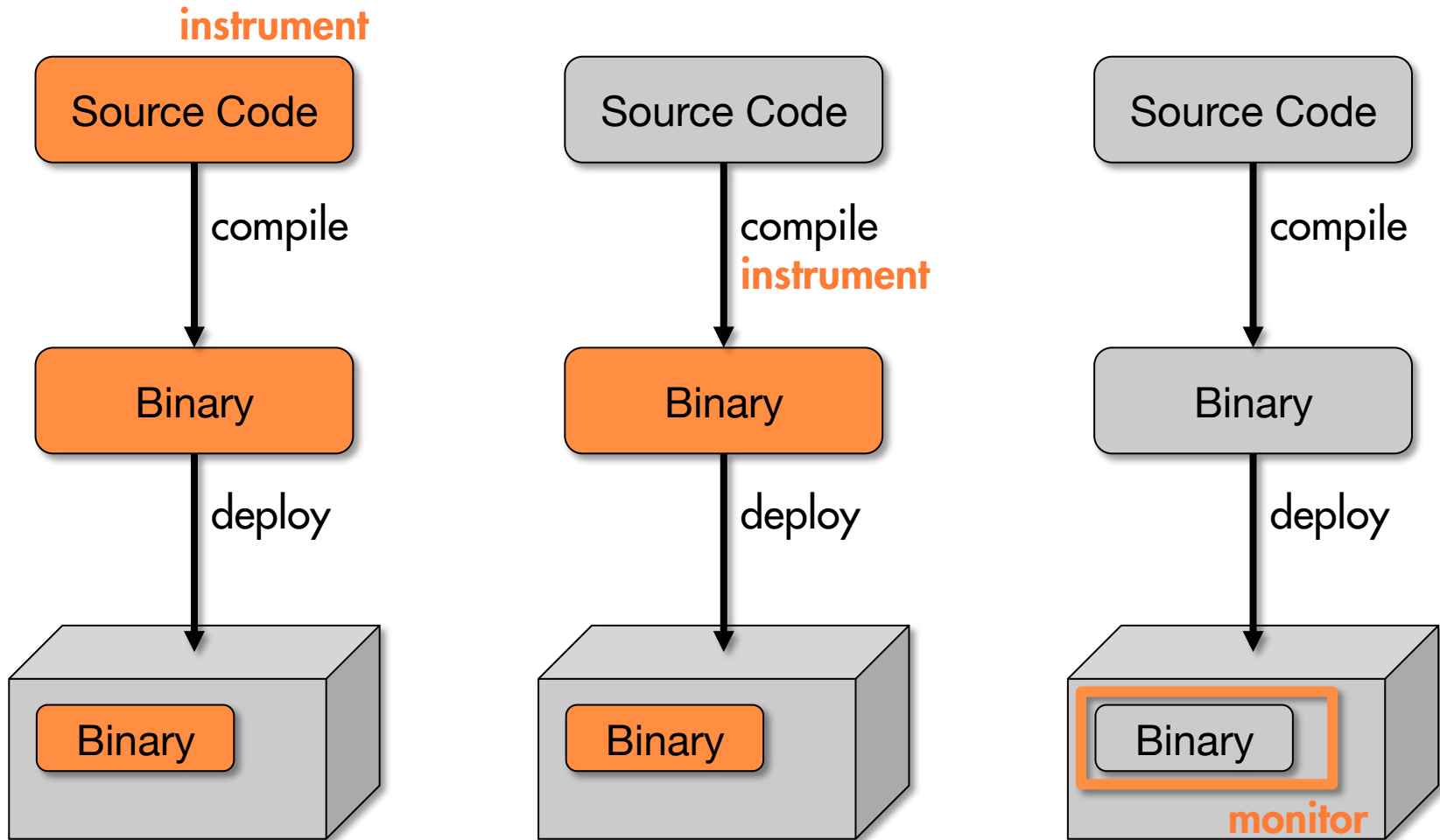
 **BULLSEYE**
TESTING TECHNOLOGY
Bullseye Coverage

 *froglogic*
Squish Coco

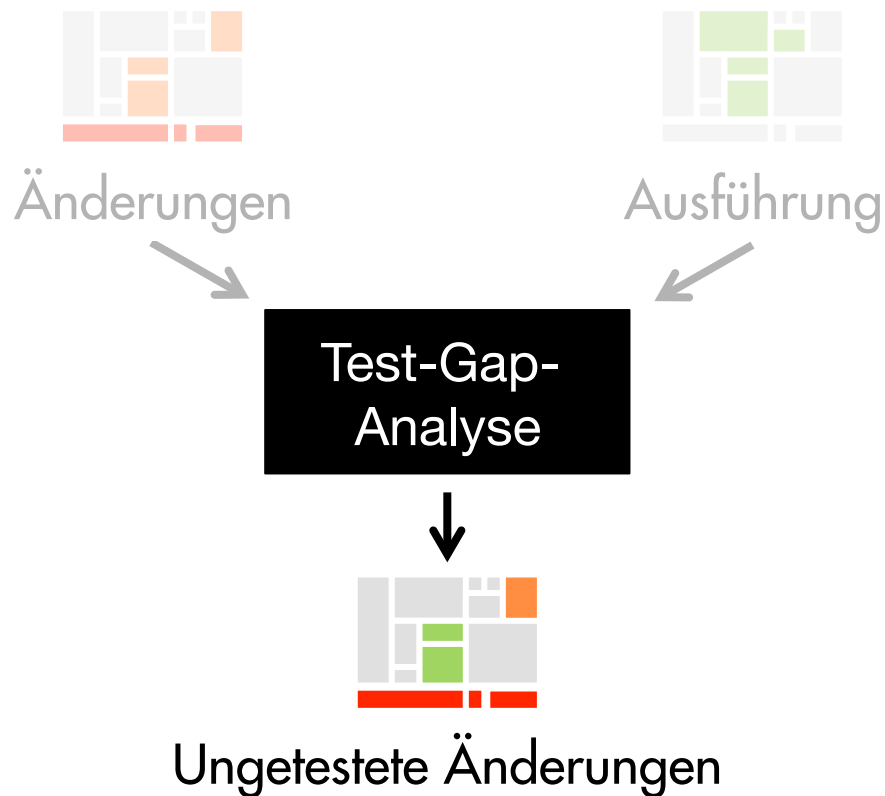
 QA SYSTEMS
The Software Quality Company
Cantata

 debug

Ansätze zur Instrumentierung



Kombination

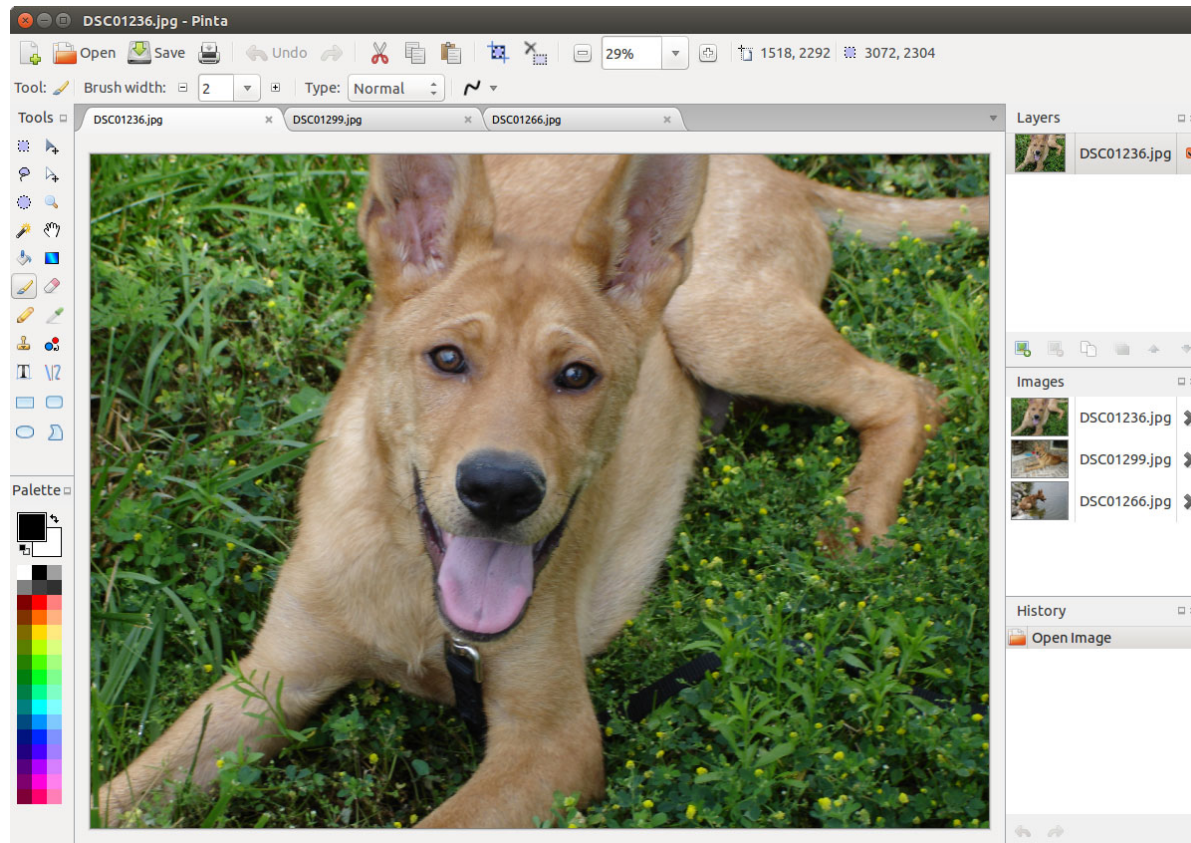


- Regelmäßig aktualisiert
- Inkrementell verzahnt
- Pro Testumgebung (z.B. Fachbereich, Unit-Tests)
- Aggregierte Sicht

Agenda

- Motivation
- Test-Gap-Analyse
- **Demo**
- Erfahrungen aus der Praxis
- Ausblick

SuT: Pinta – Ein .NET Open Source Zeichenprogramm



DEMO



www.teamscale.com



100% Change Coverage

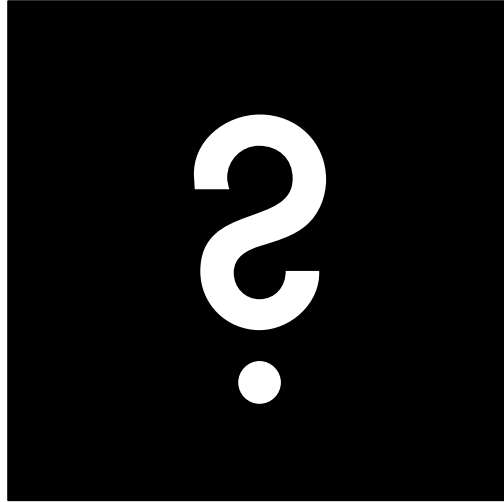
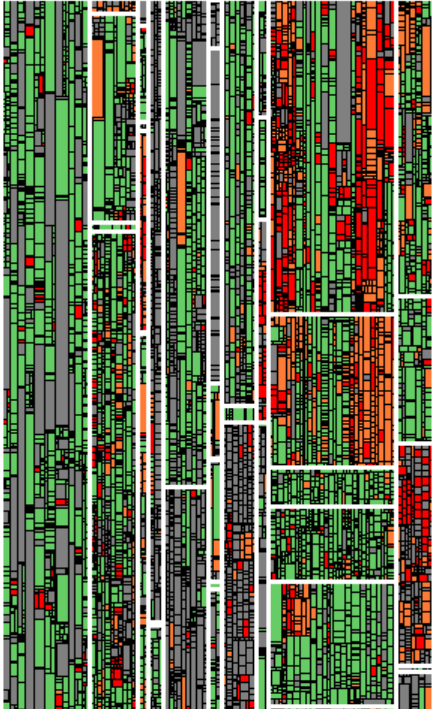


100% Change Coverage → 0 Fehler

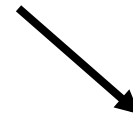
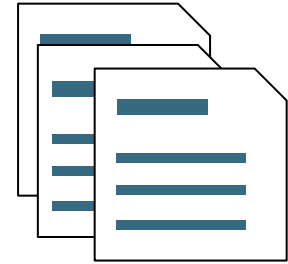
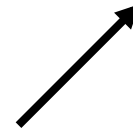
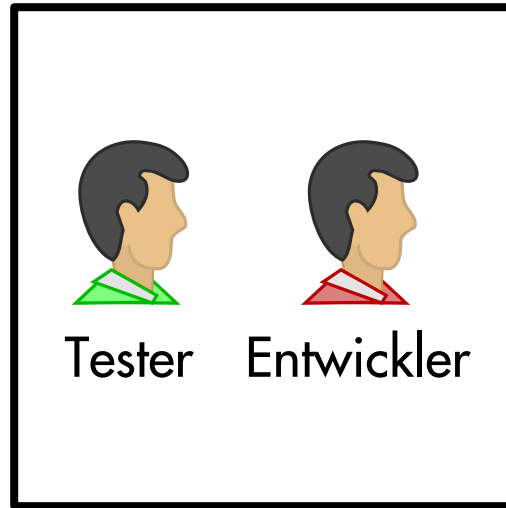
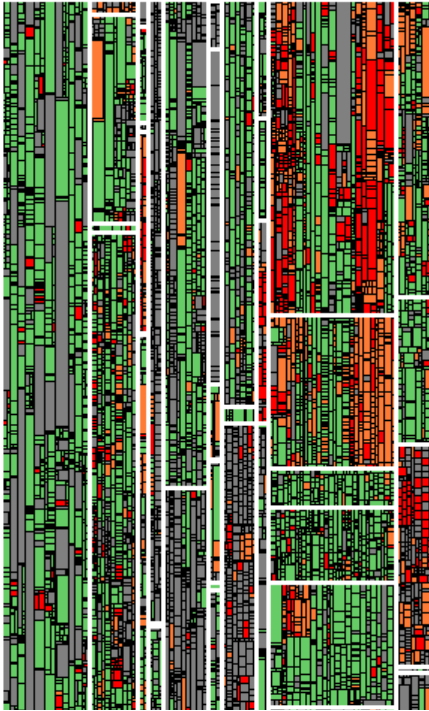
Agenda

- Motivation
- Test-Gap-Analyse
- Demo
- **Erfahrungen aus der Praxis**
- Ausblick

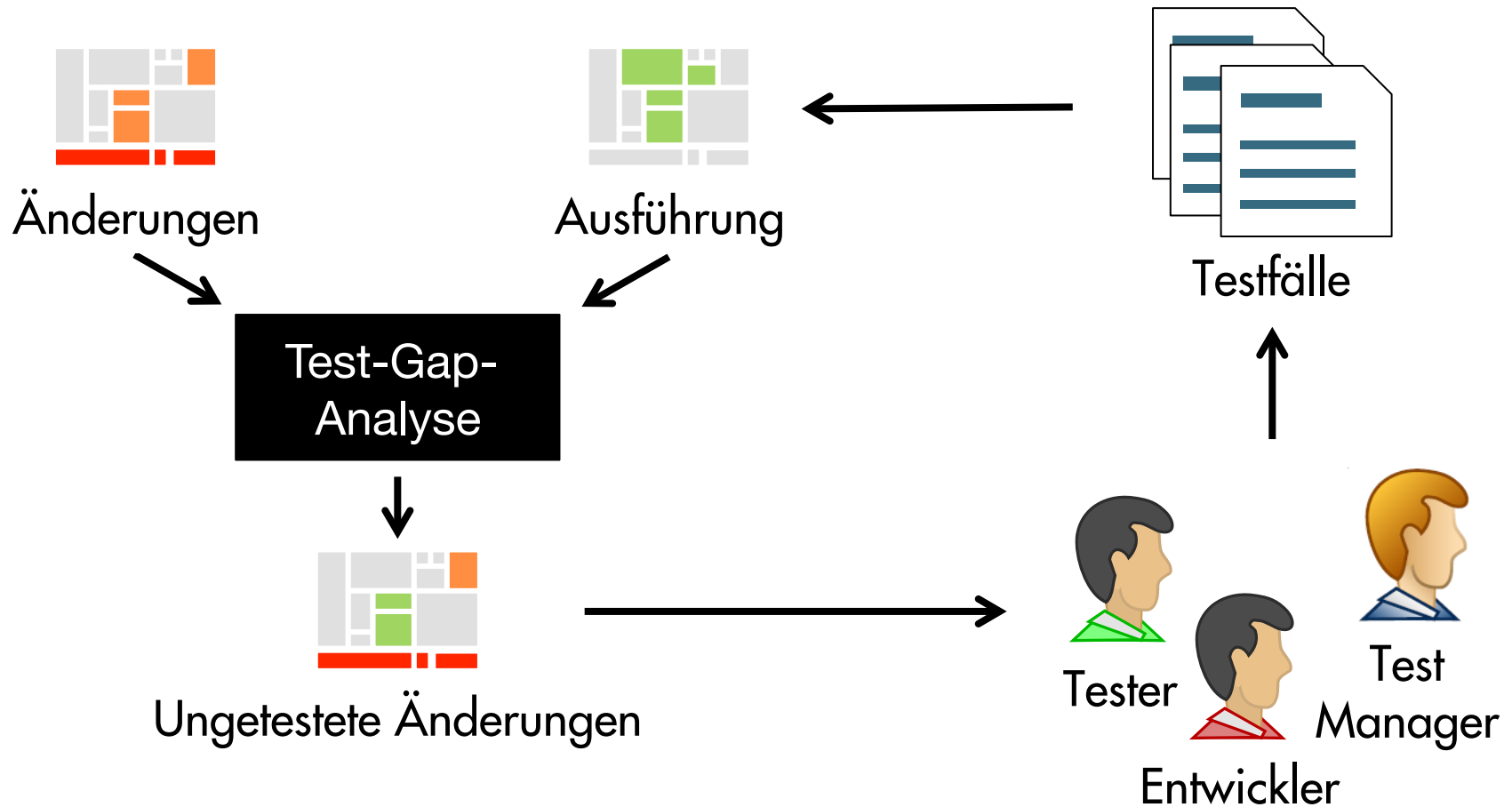
Testauswahl



Testauswahl



Einsatz im Testprozess



Test-Arten

Hotfix-Test

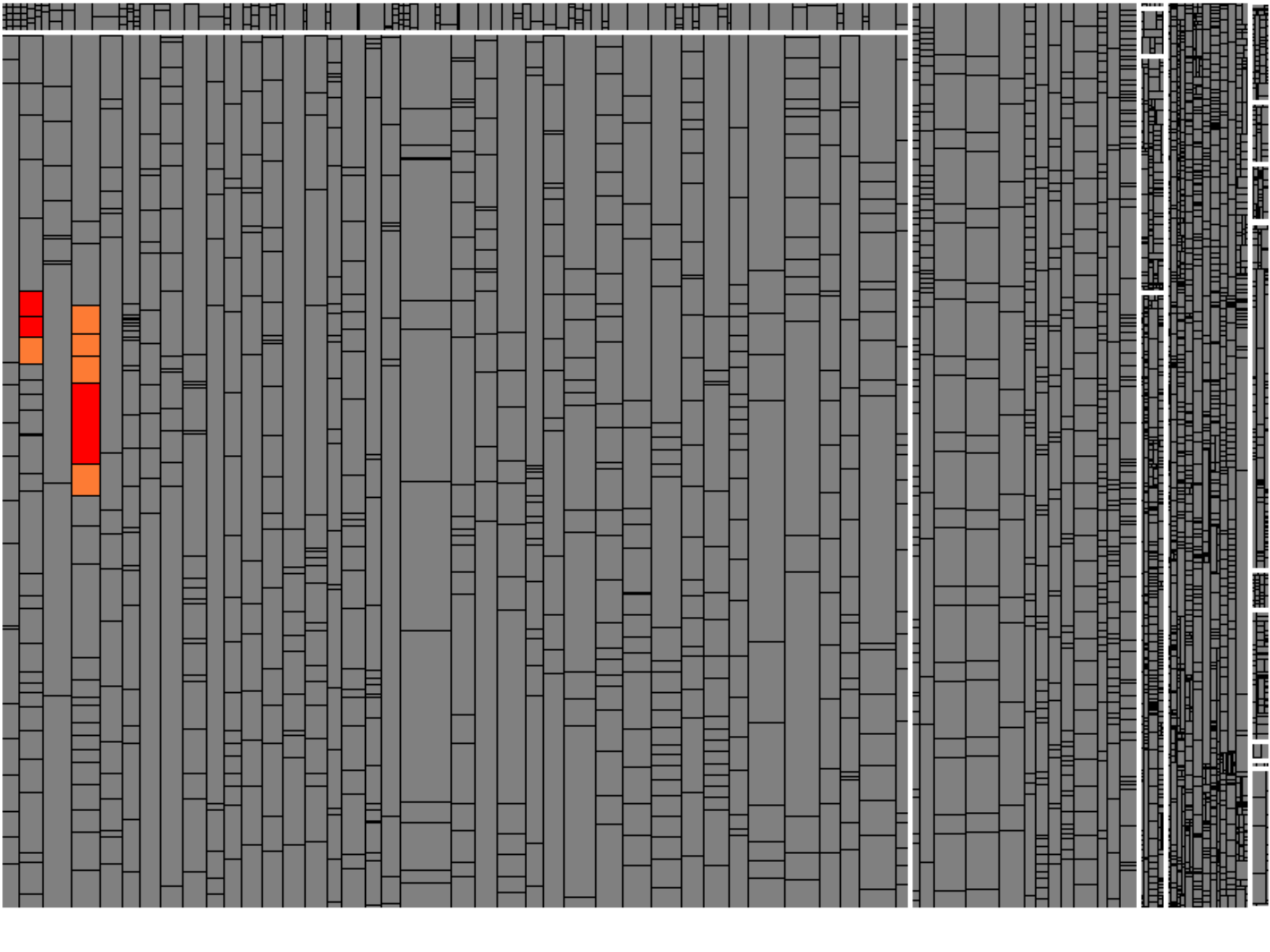
- On-Demand, wenig Zeit
- Getestet werden nur Änderungen des Hotfix

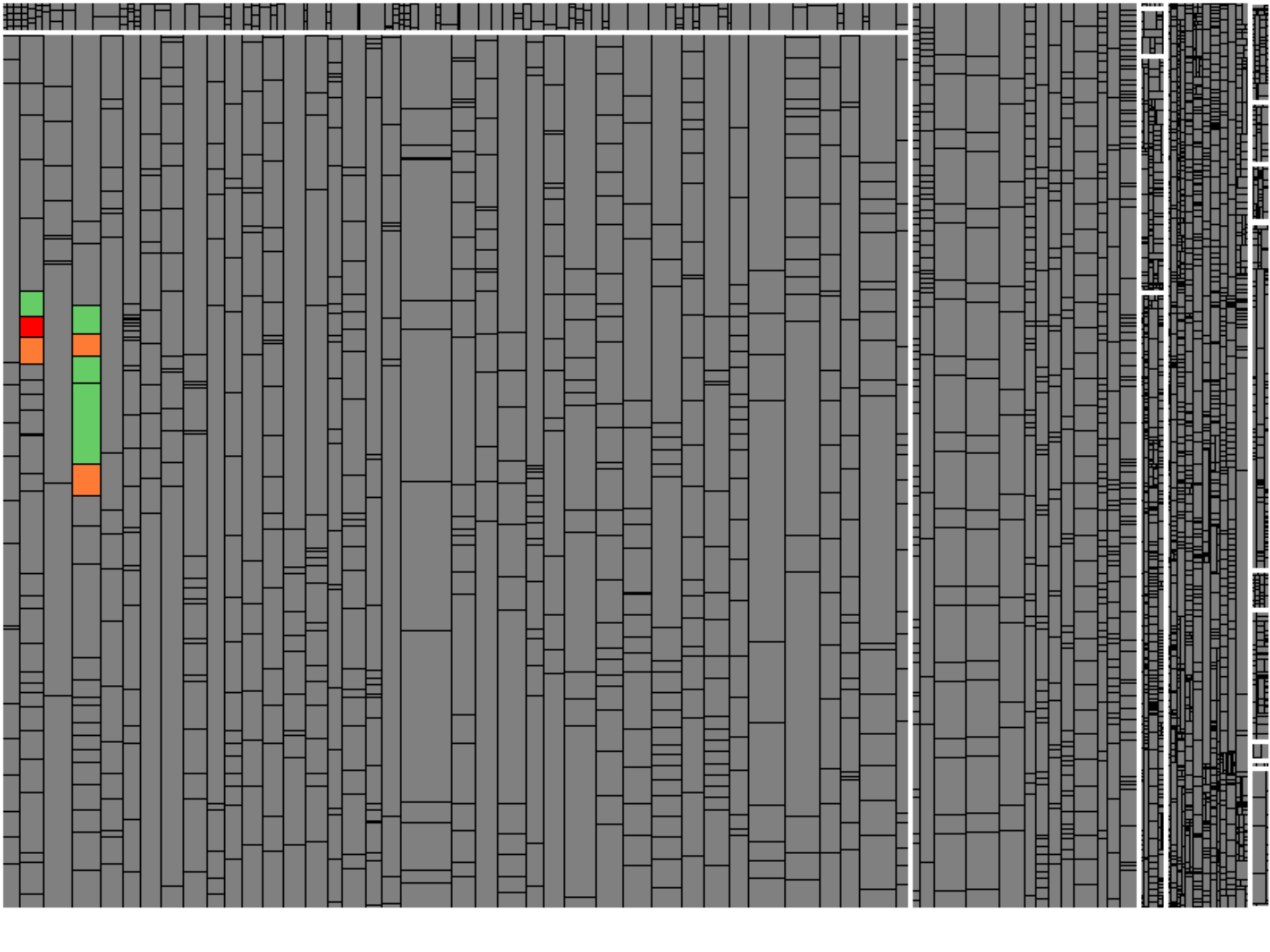
Iterationstest

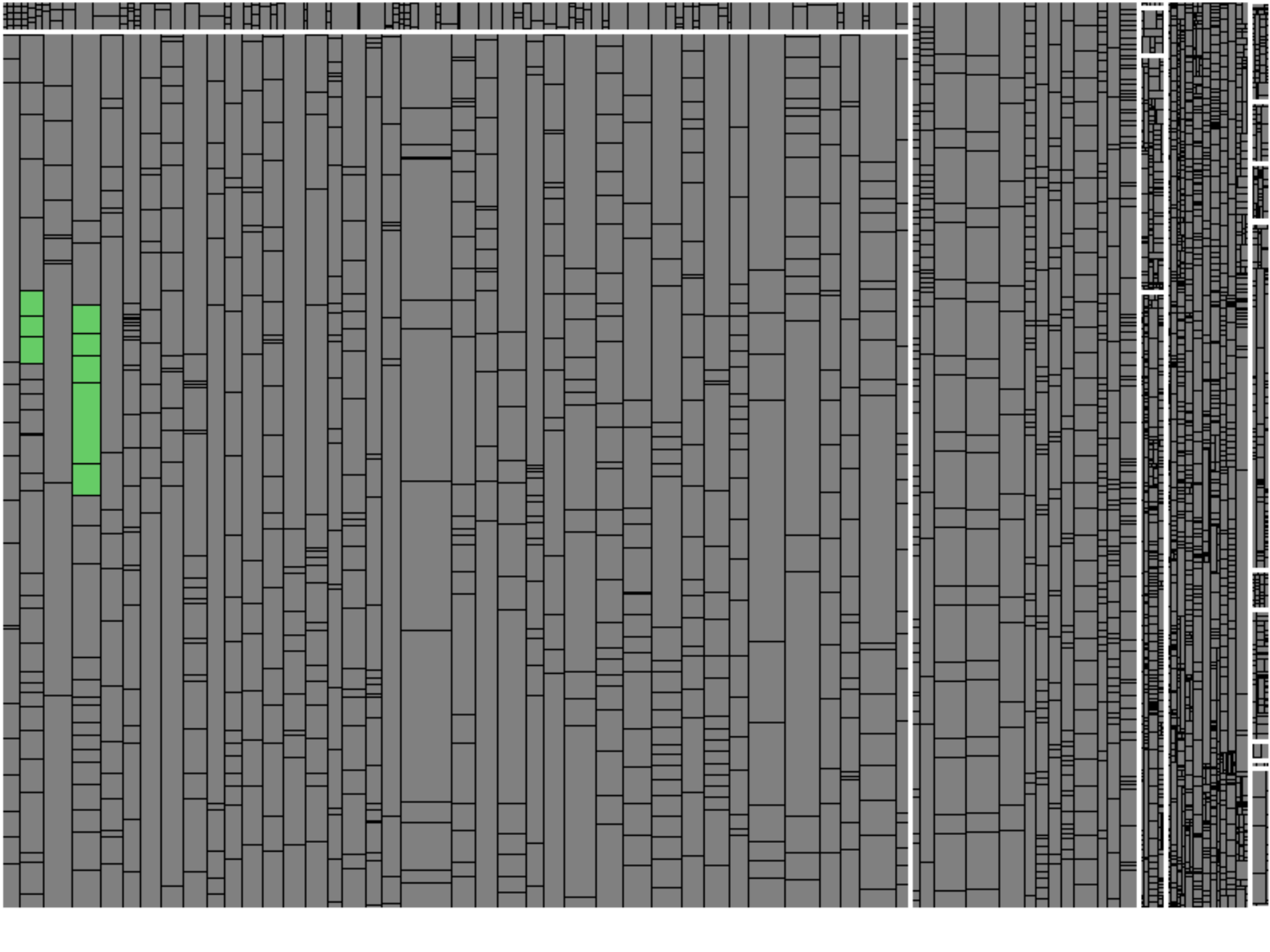
- Entwicklungsbegleitend, kontinuierlich, Iterationsdauer (2-4 Wochen)
- Getestet werden Änderungen aus der Iteration

Release-Test

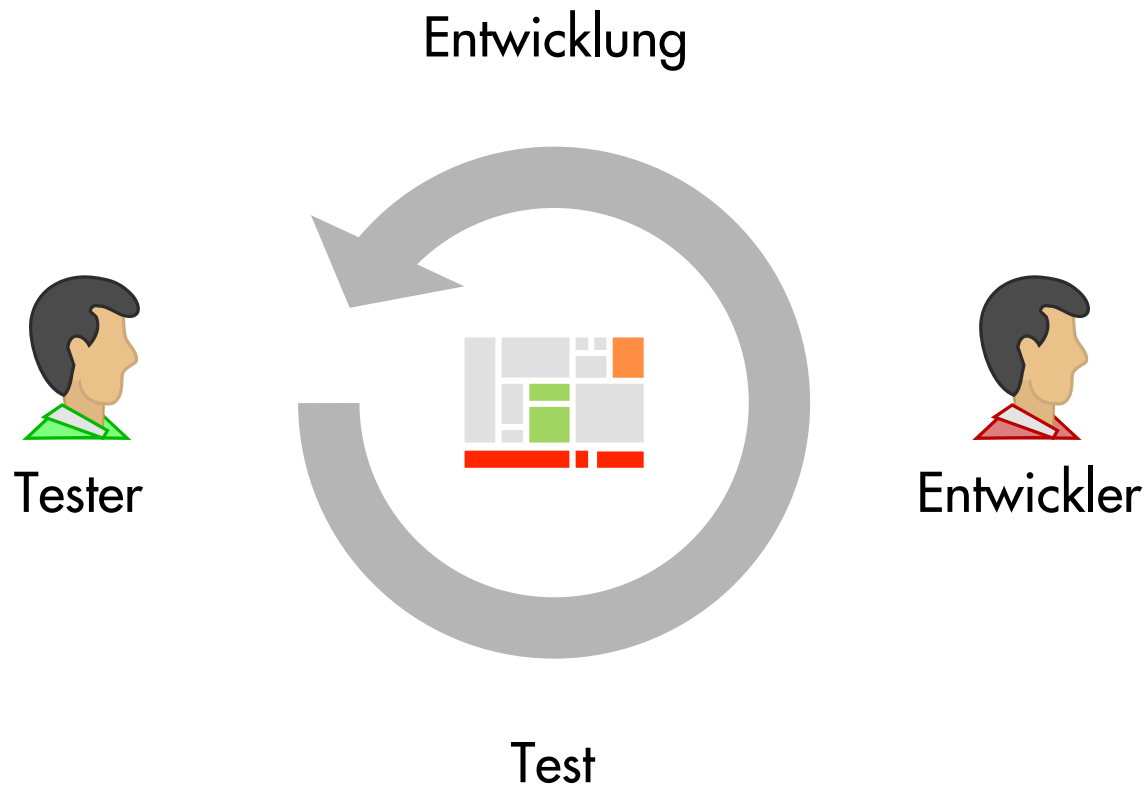
- Vor Release, Dauer mehrere Wochen
- Getestet werden alle Änderungen seit letztem Release
- 2-4 Mal pro Jahr, je nach Release-Zyklus

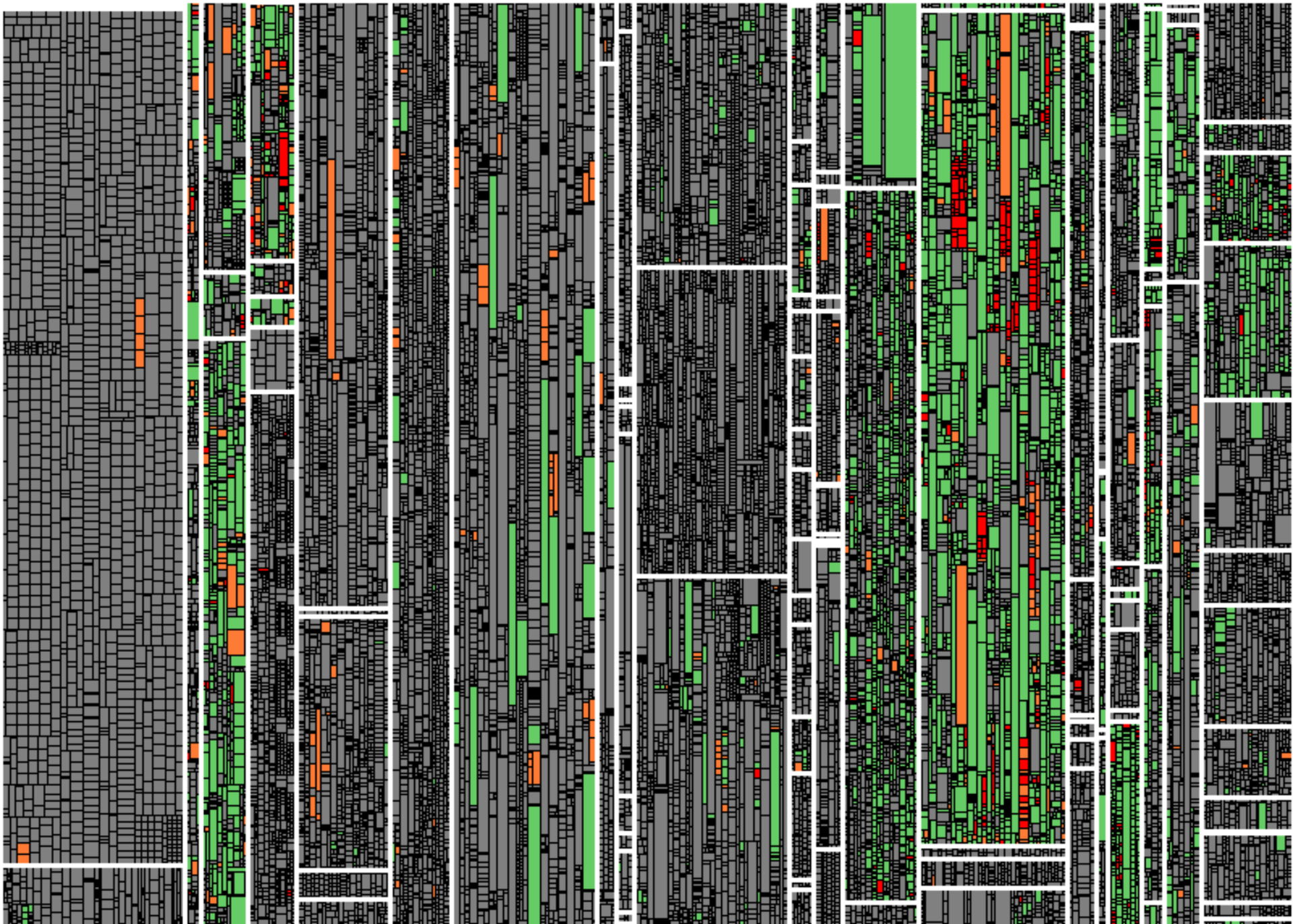


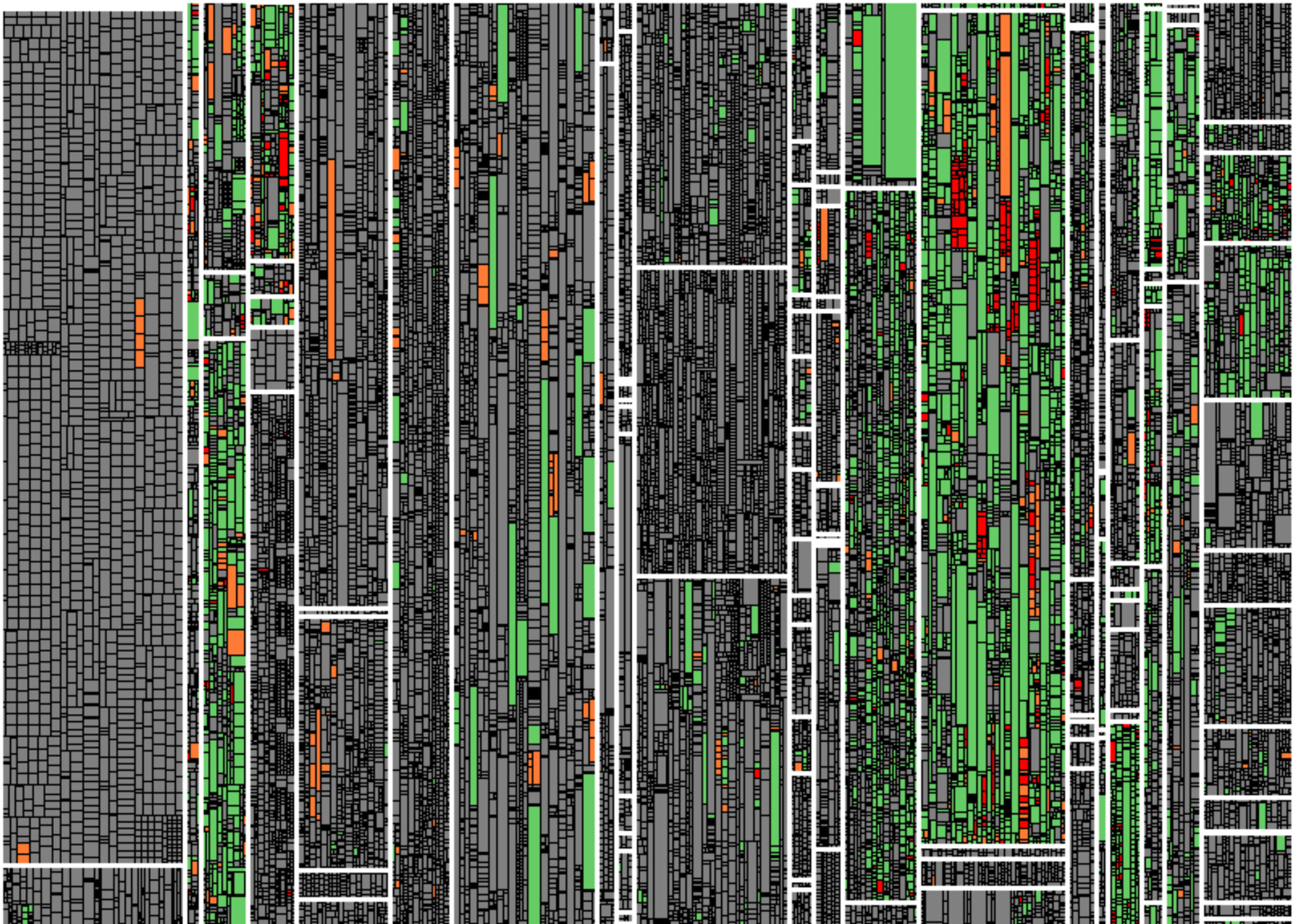


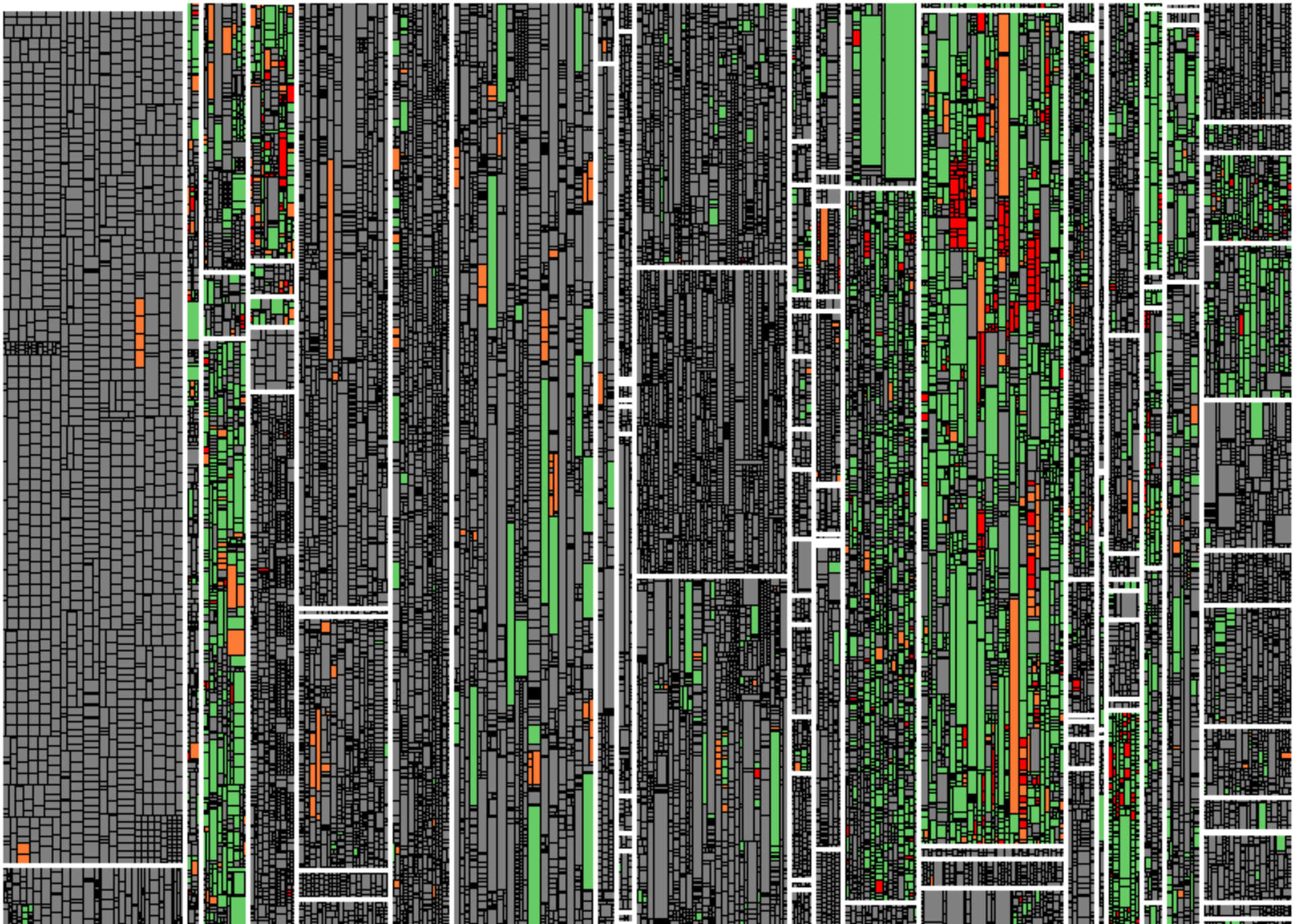


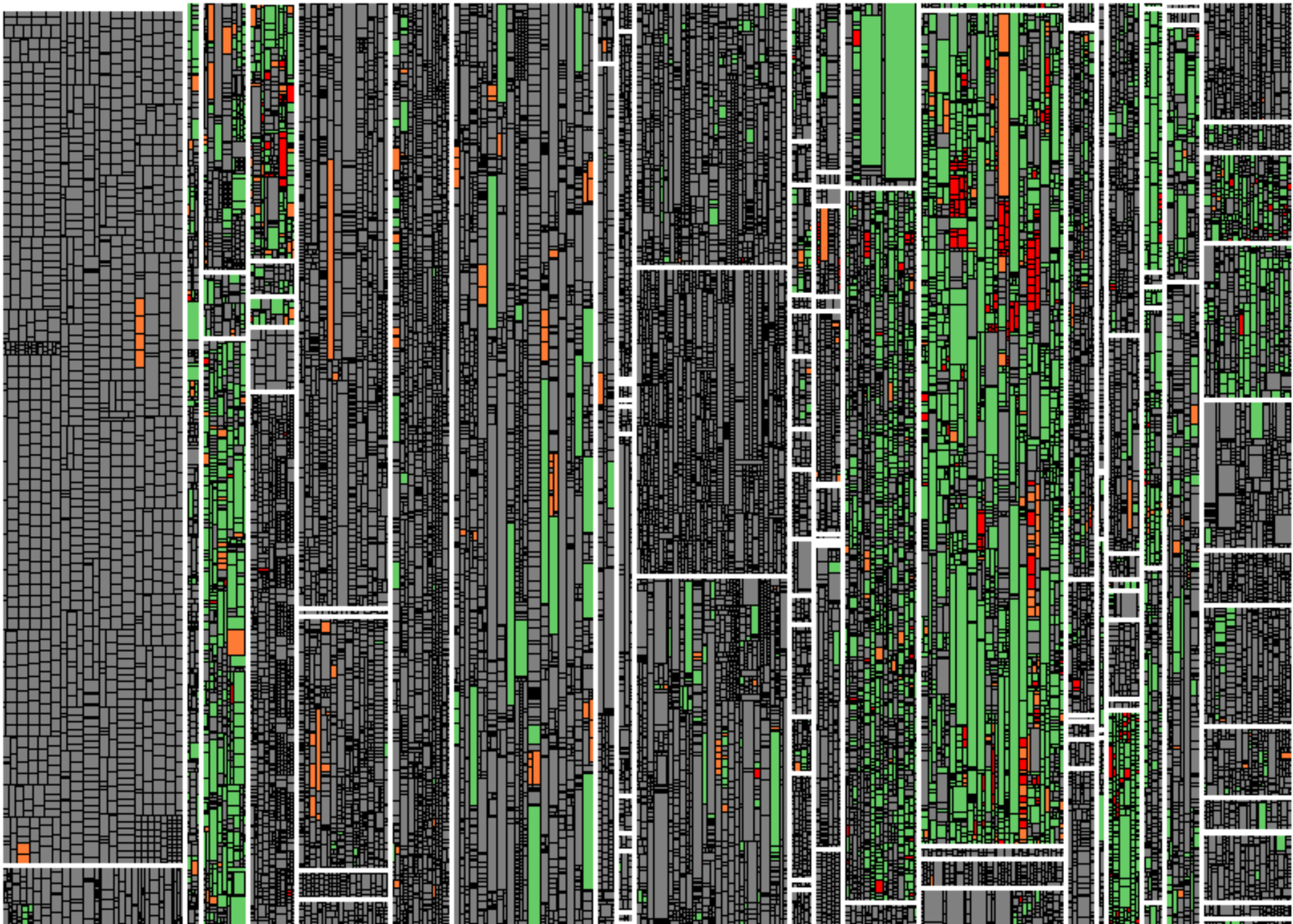
Iterationstest

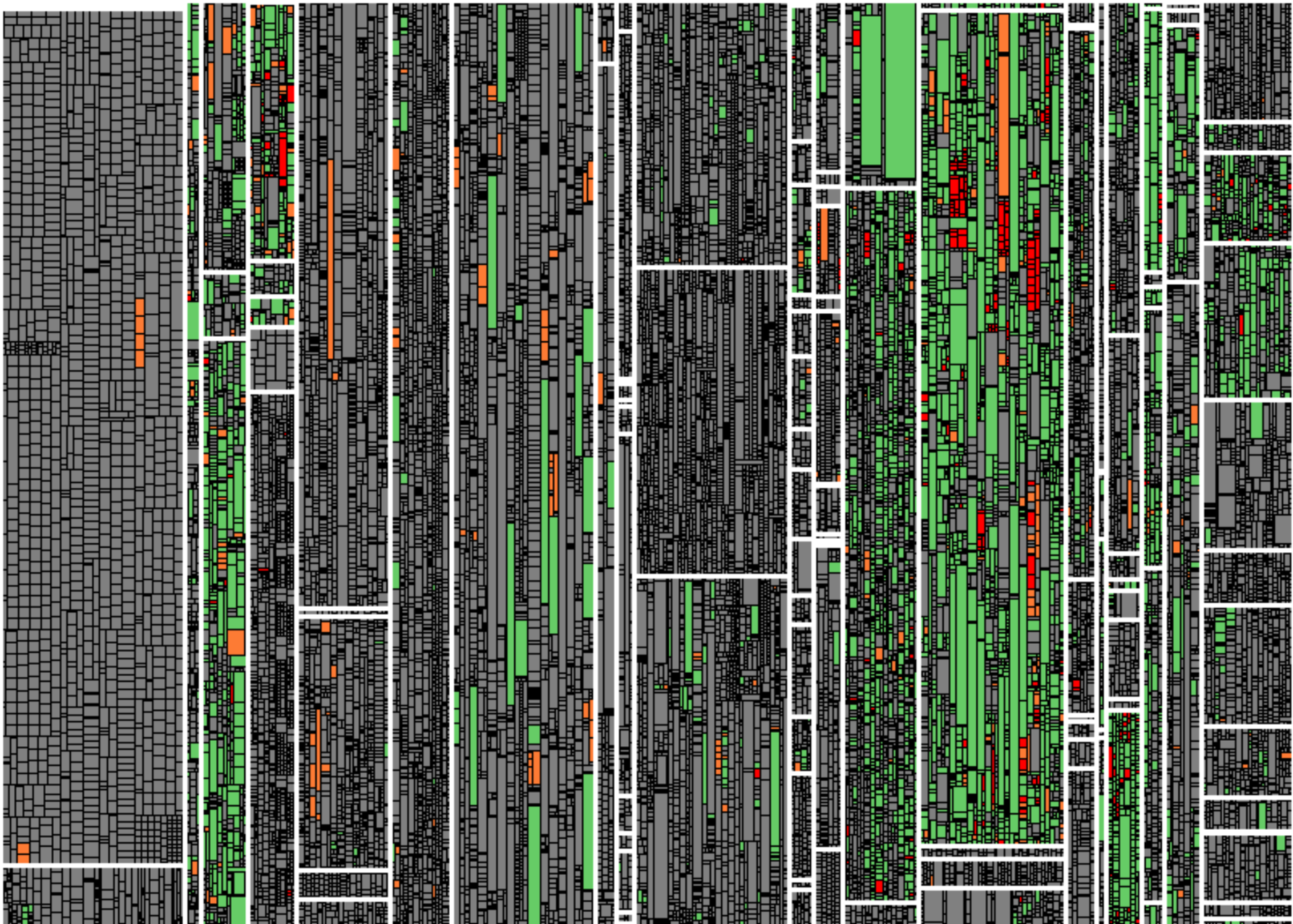


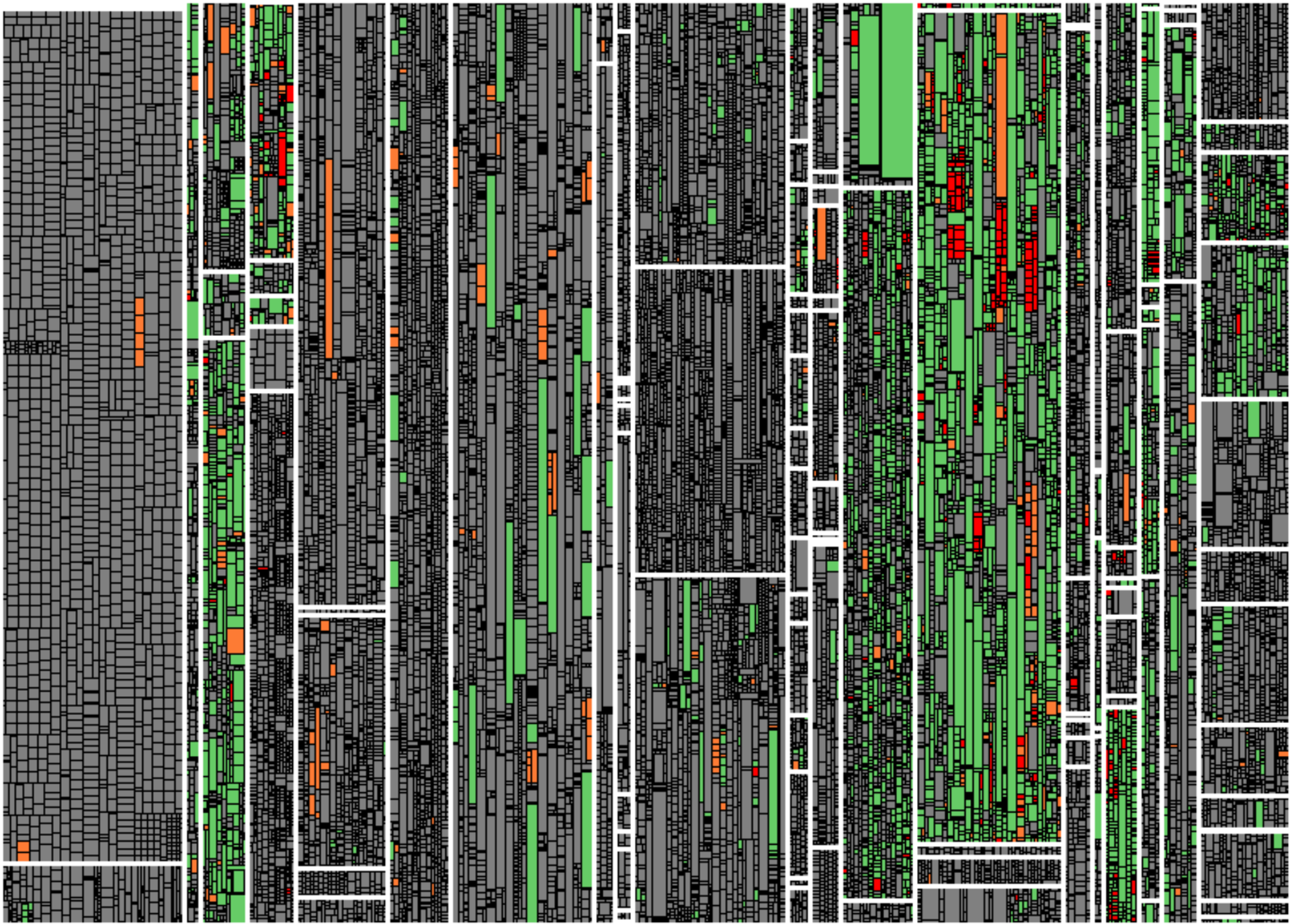


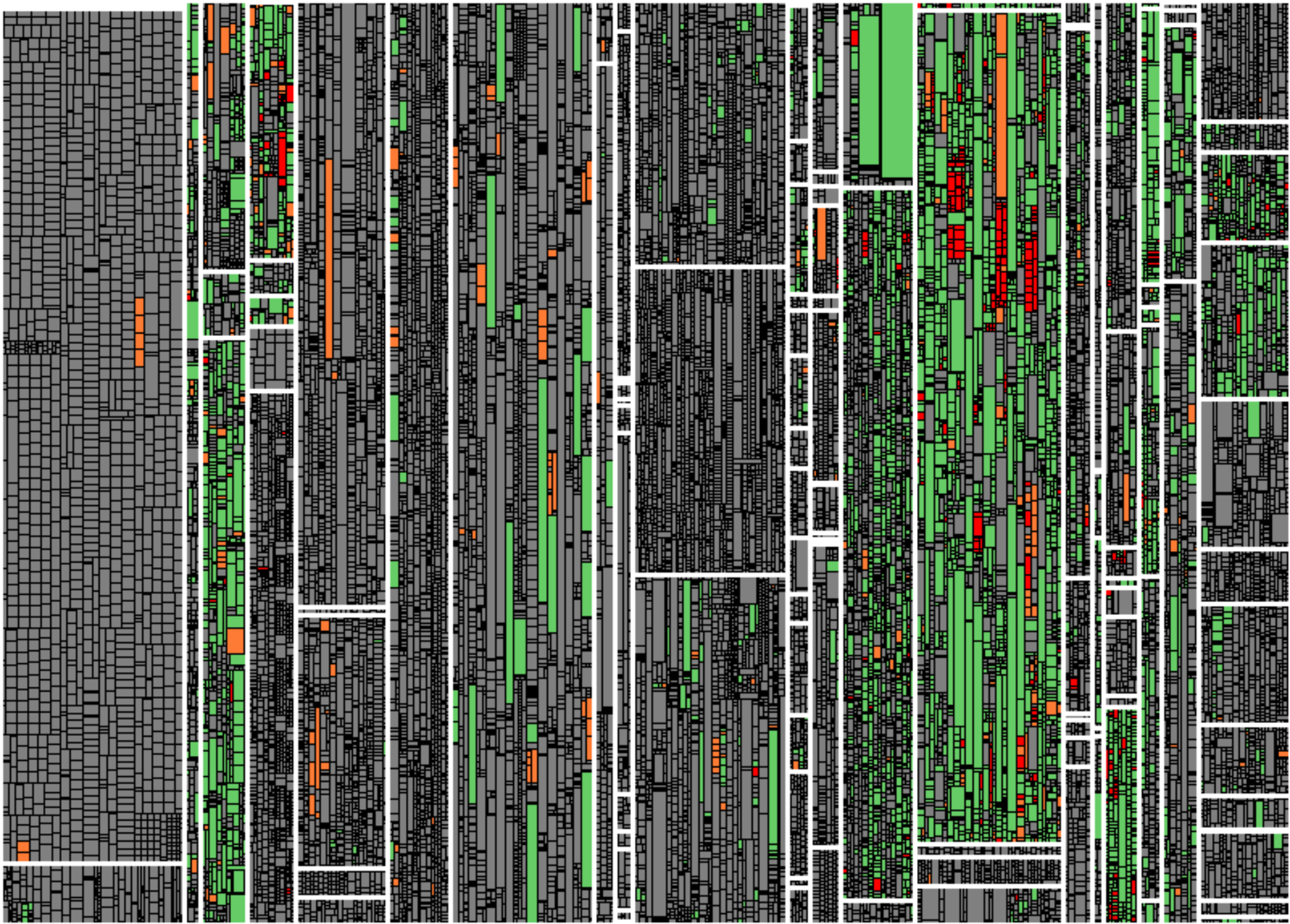


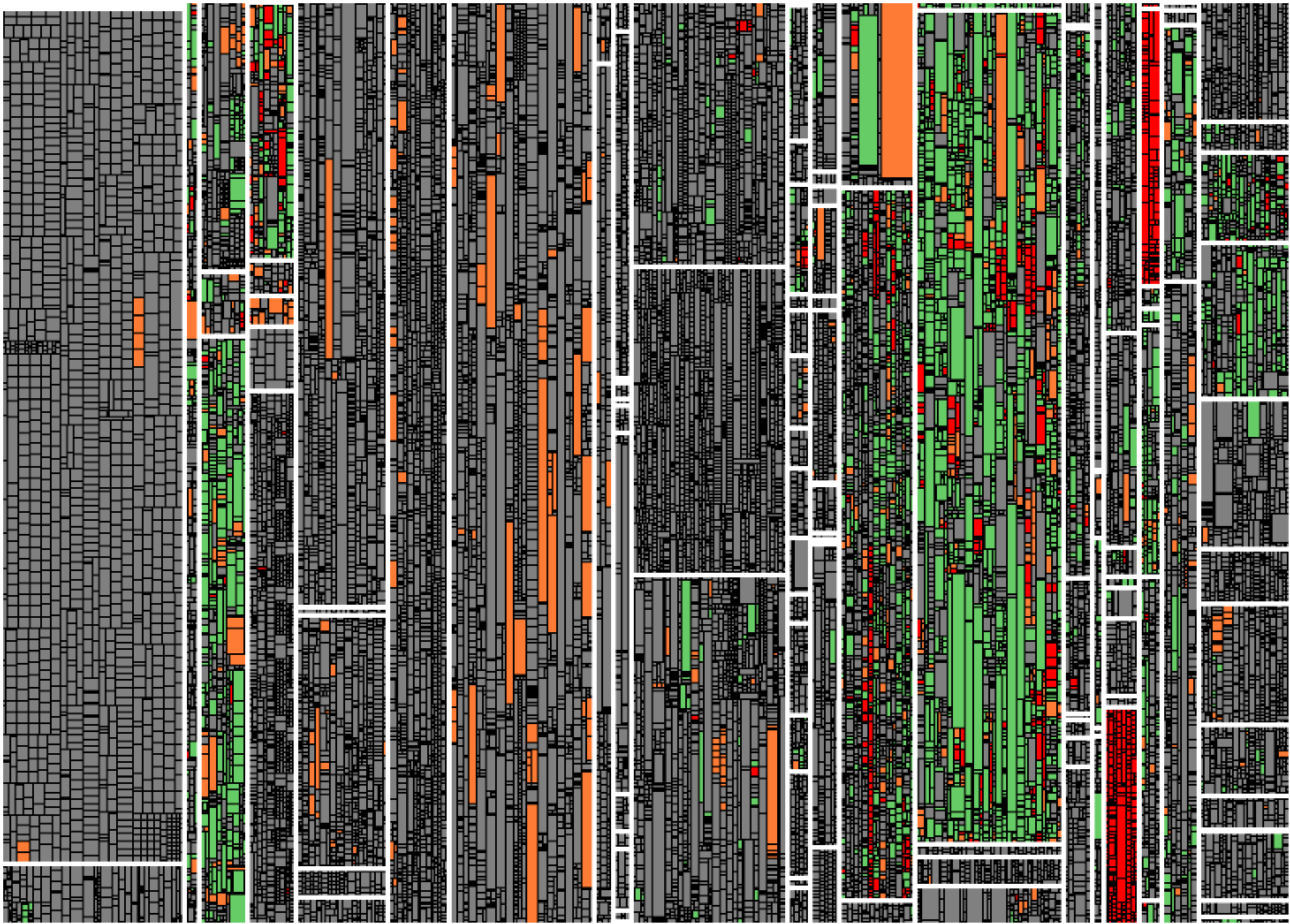


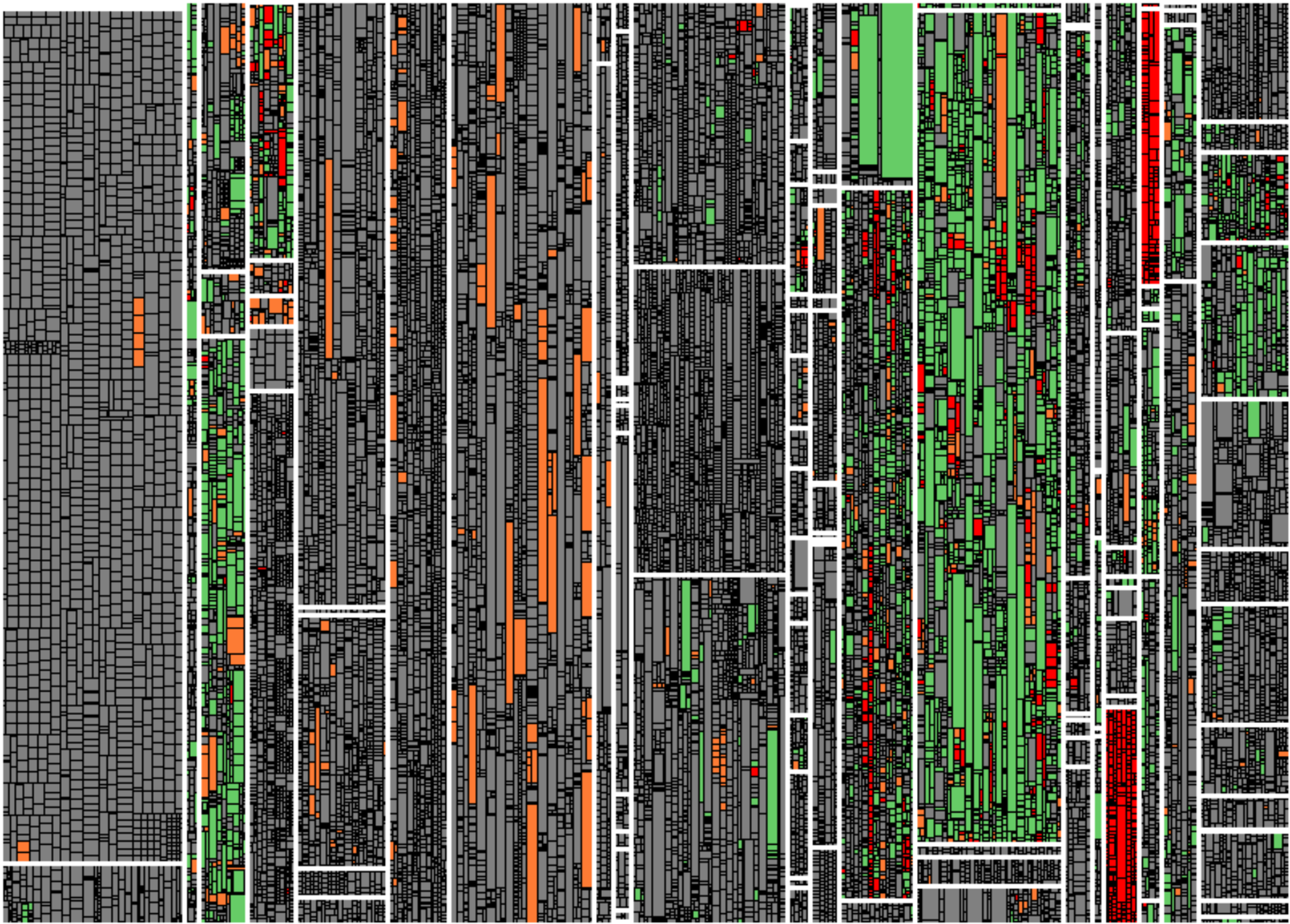


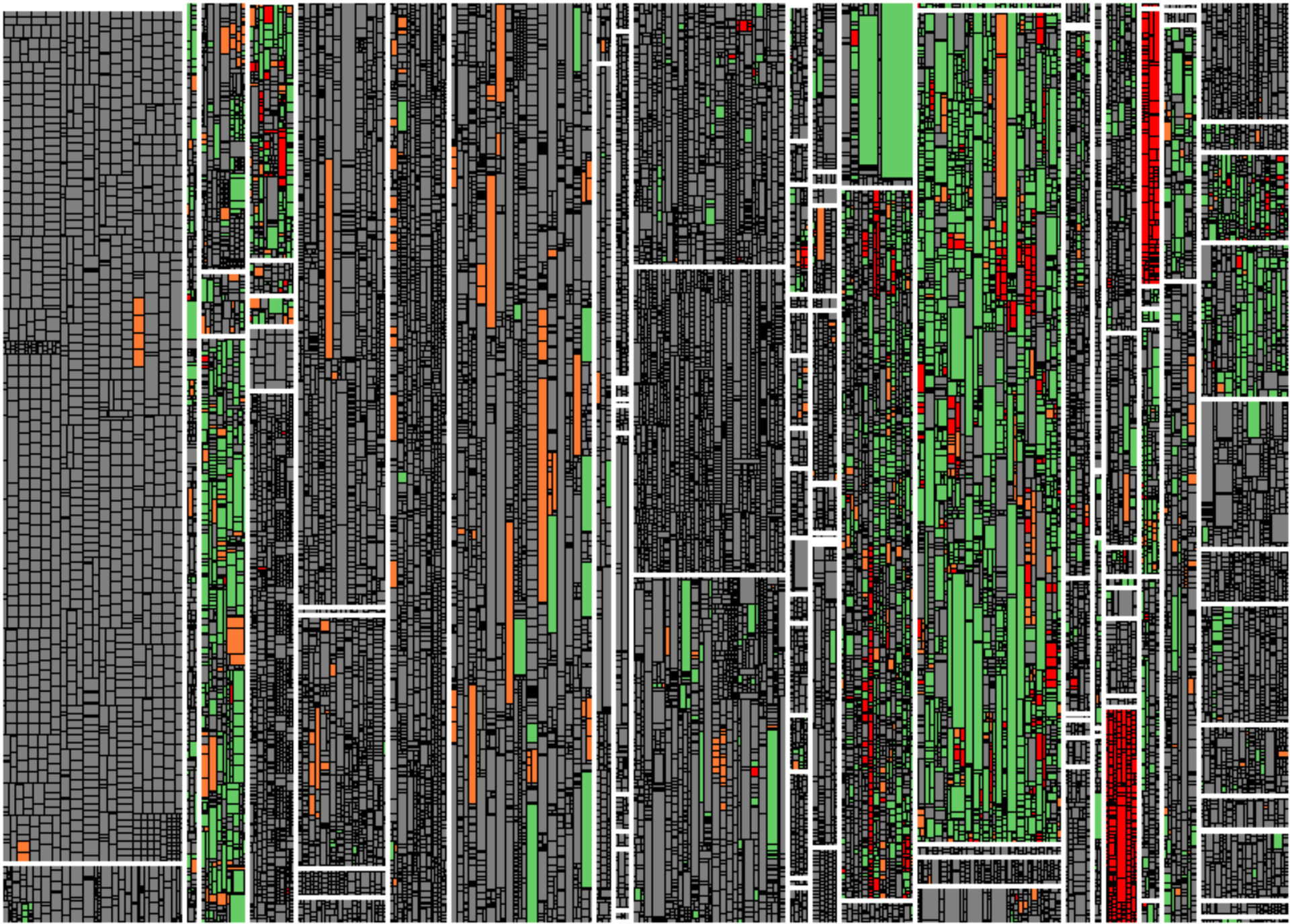


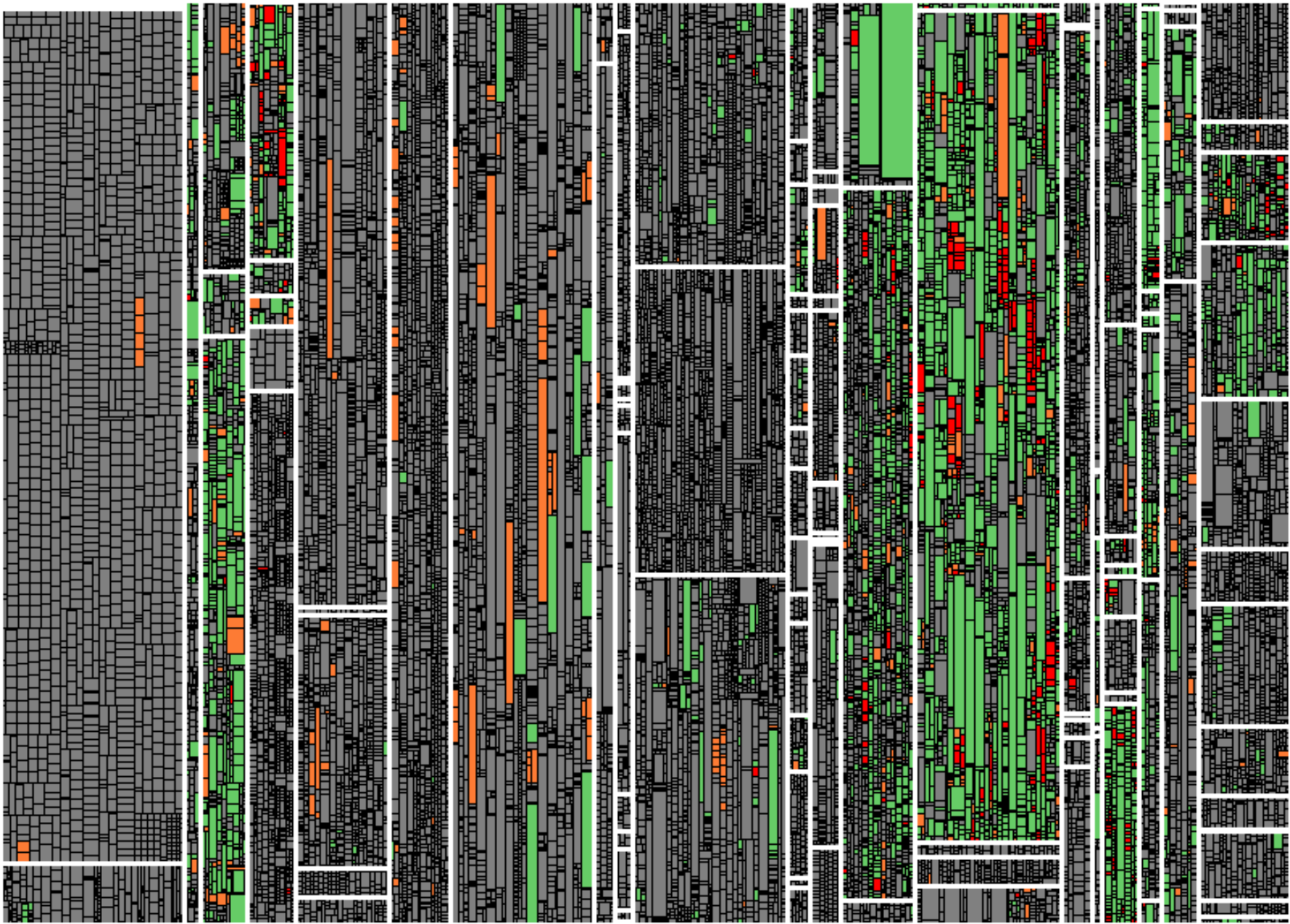


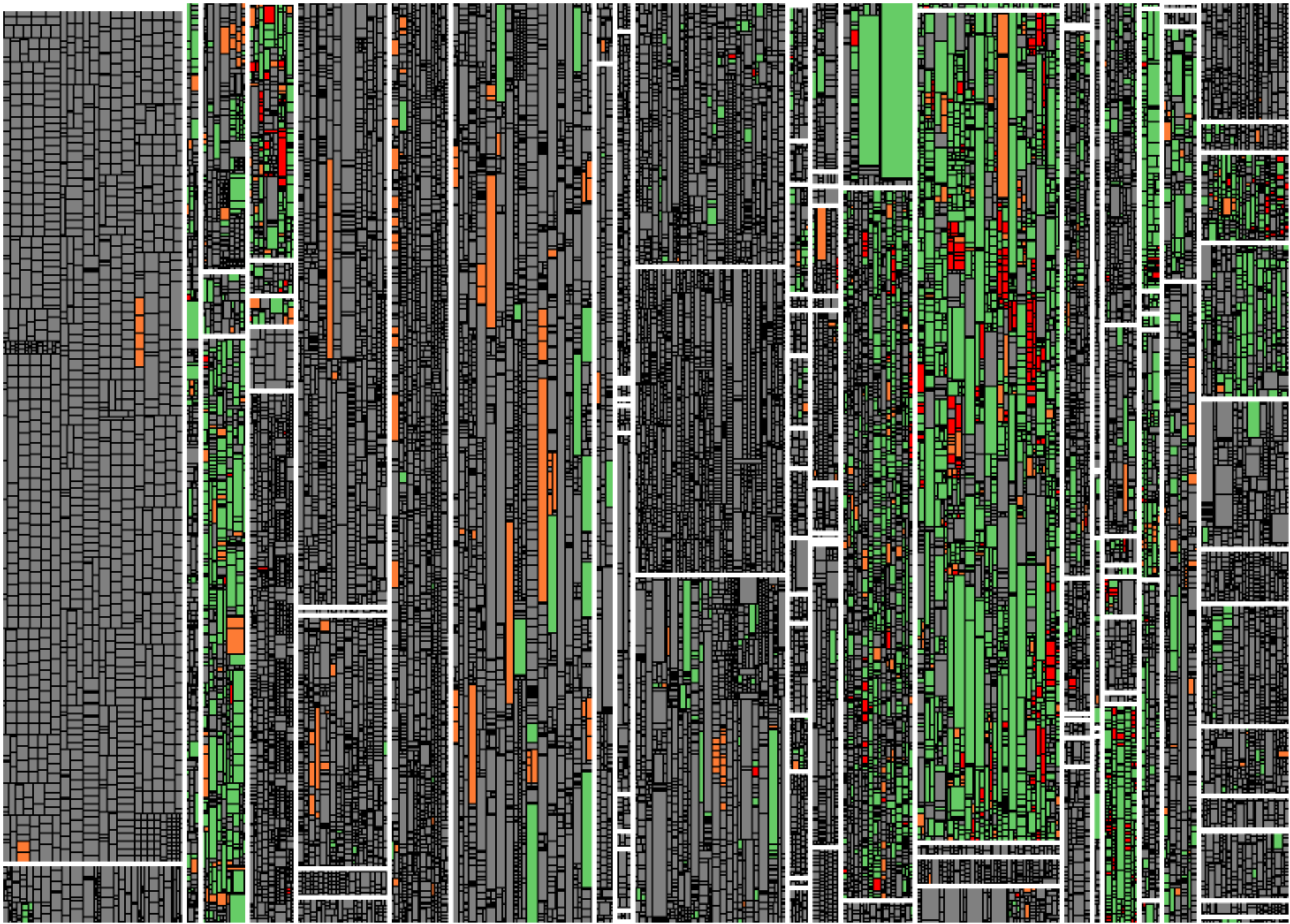


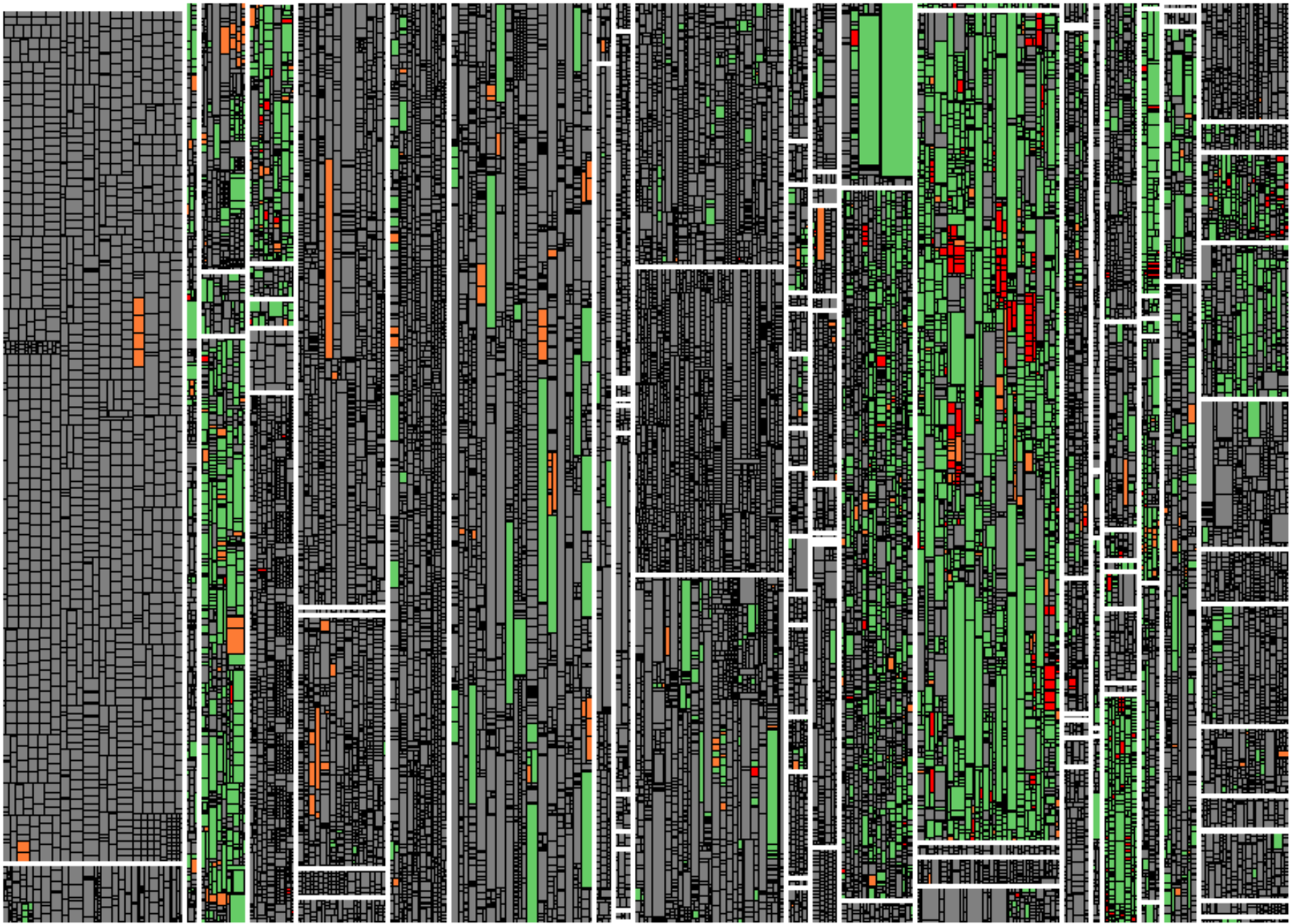


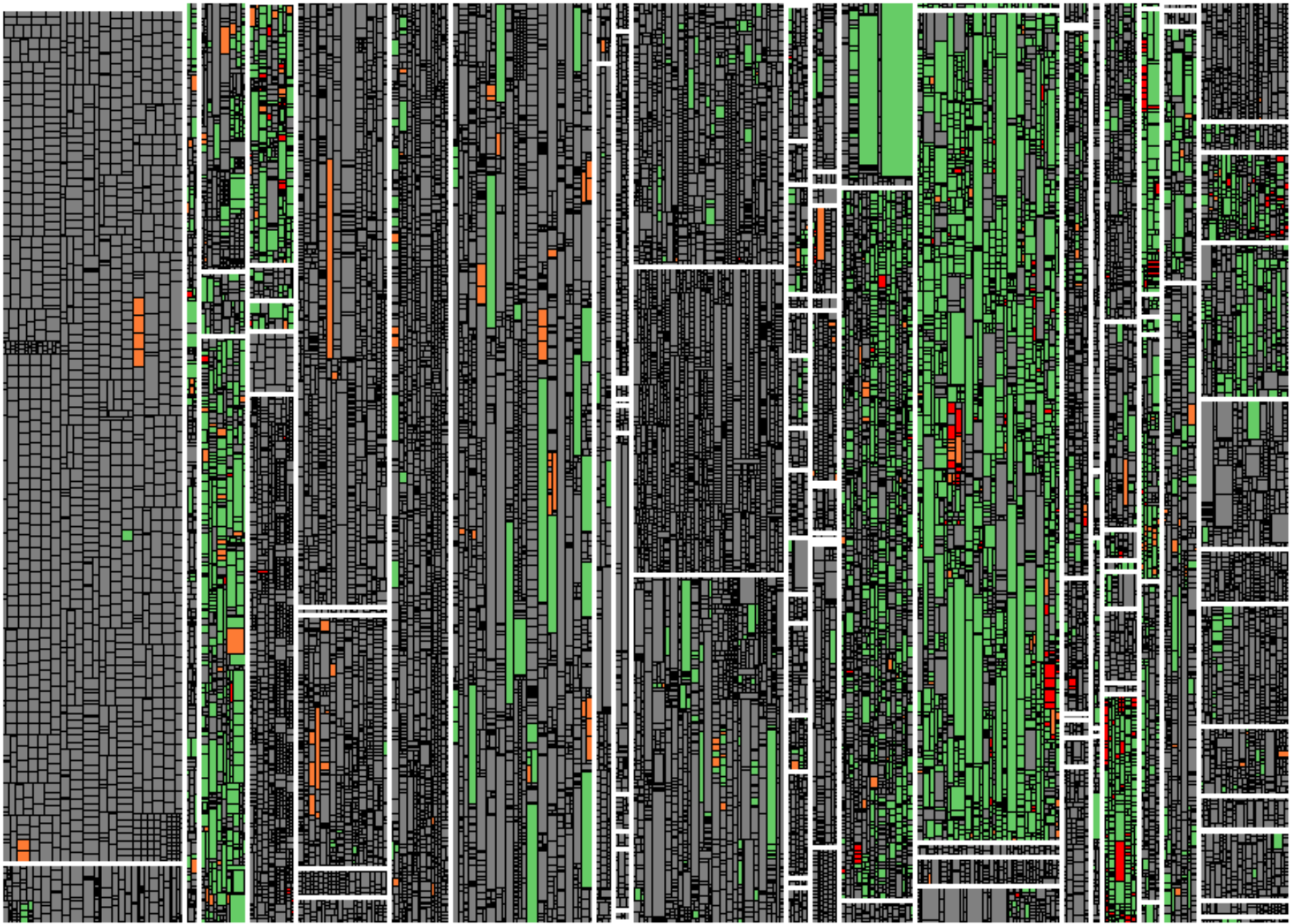






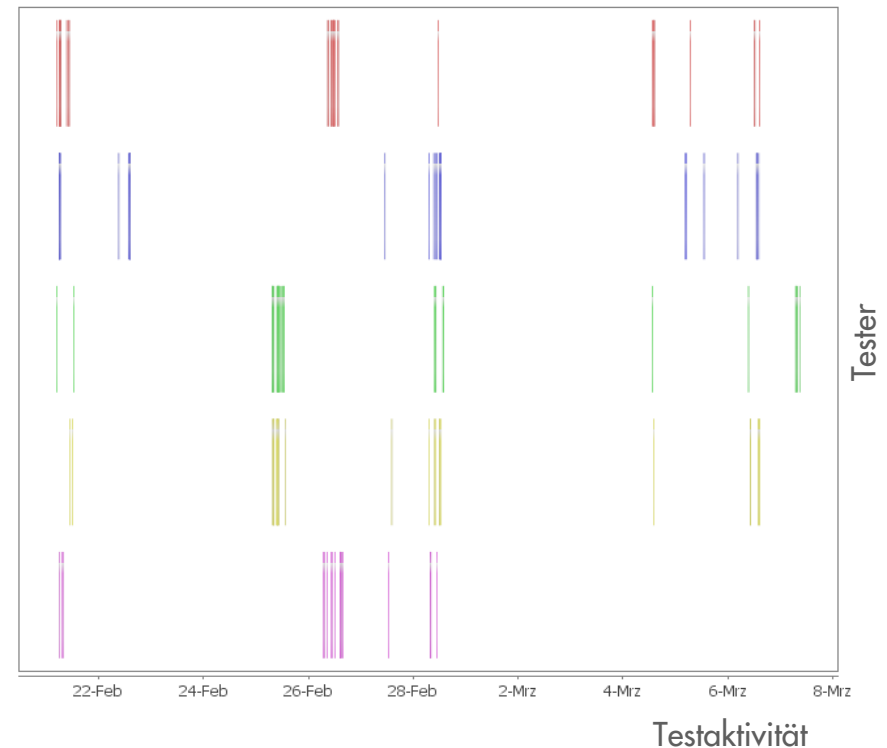


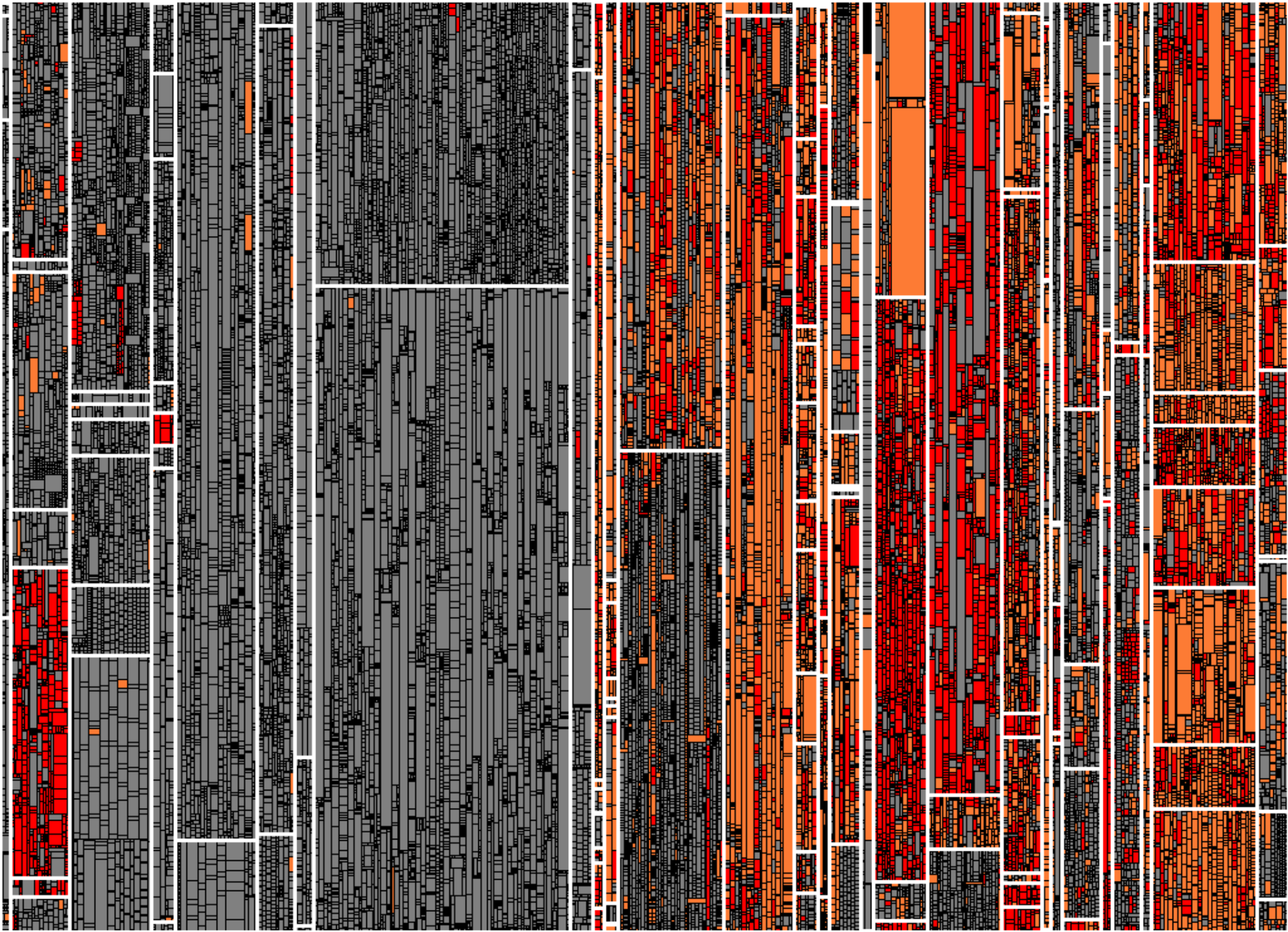


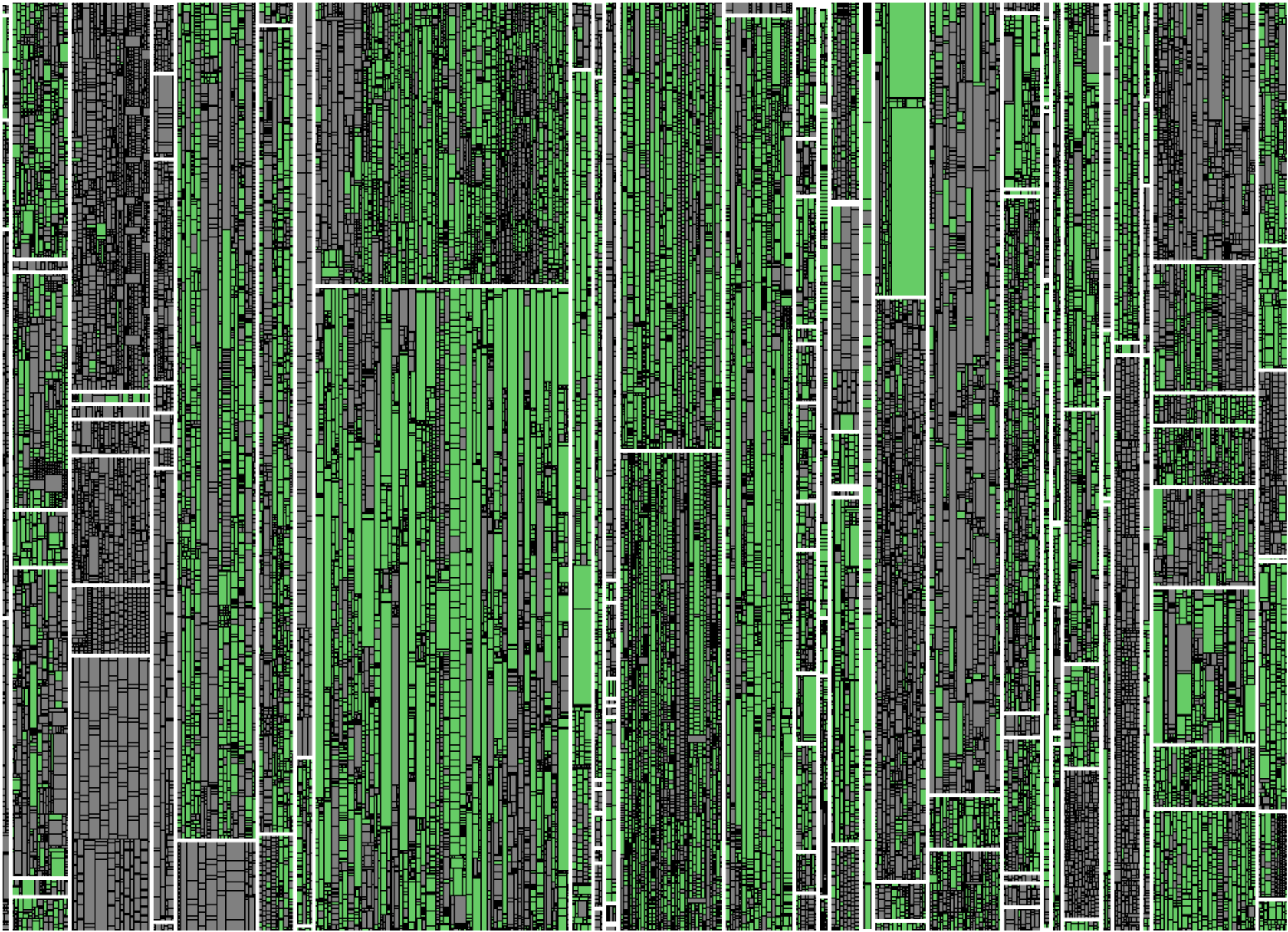


Release-Test: Test-Factory

- 5 Tester in Indien
- Regressionstest: 2 Wochen
- 143 Test-Sessions (manuell)
- 6 Systemversionen

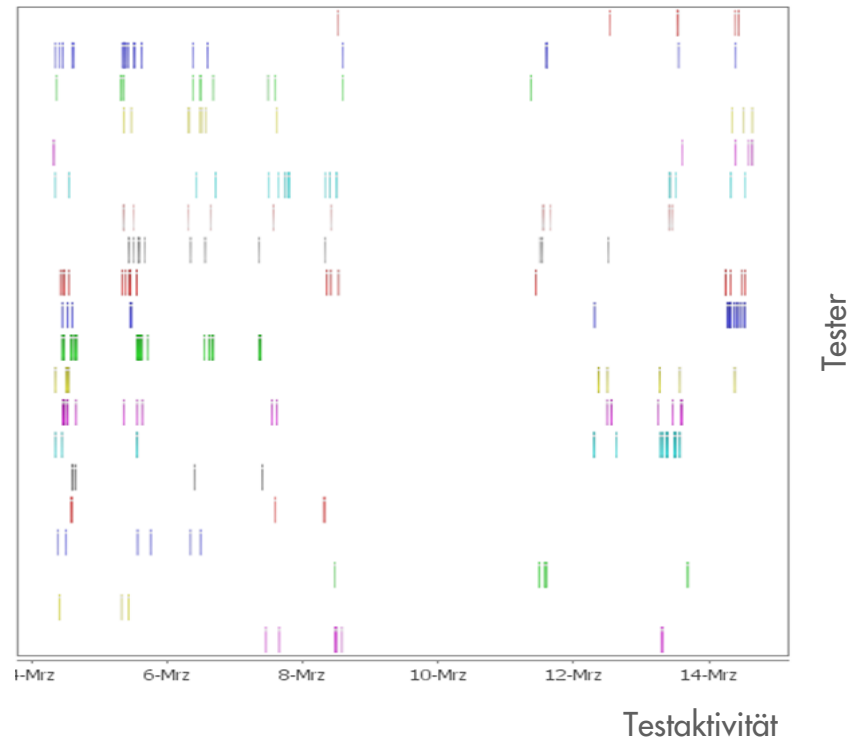


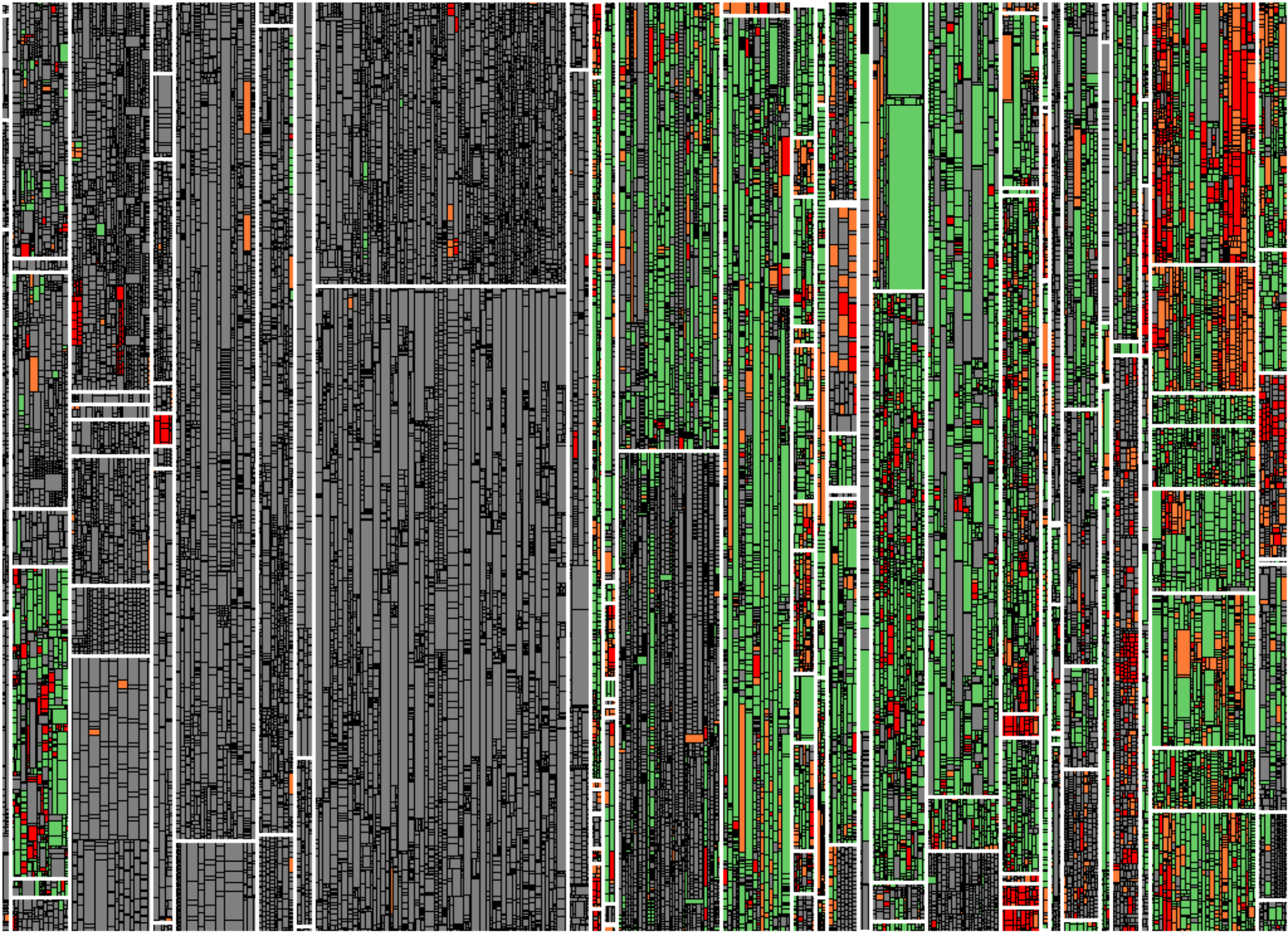




Release-Test: Explorativer Test

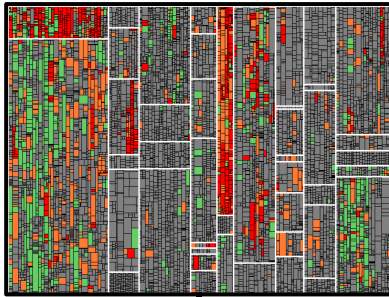
- 23 Tester
 - Requirements Engineers
 - Fachbereich
 - Endnutzer
- 2 Wochen
- 307 Test-Sessions
- 8 Systemversionen



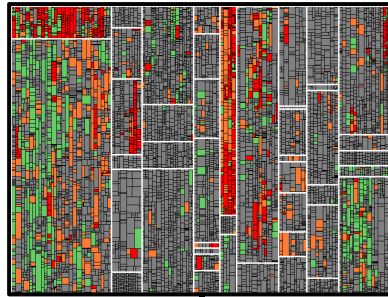


Release-Test: Verschiedene Umgebungen

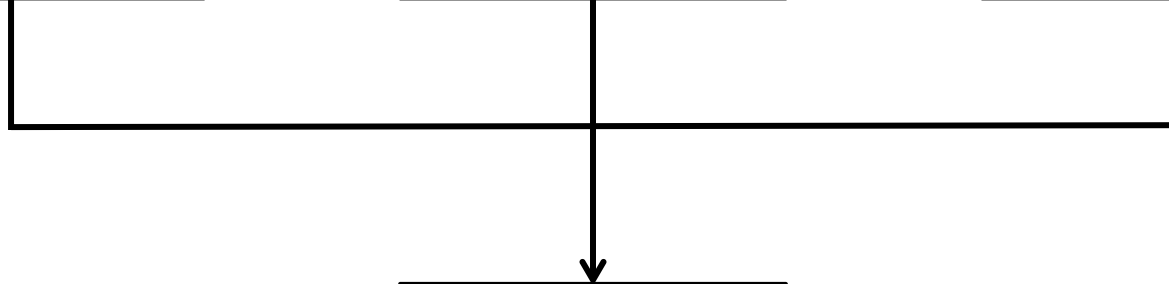
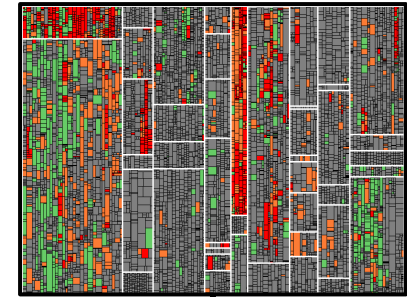
Test



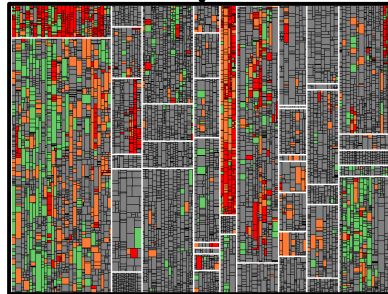
Dev



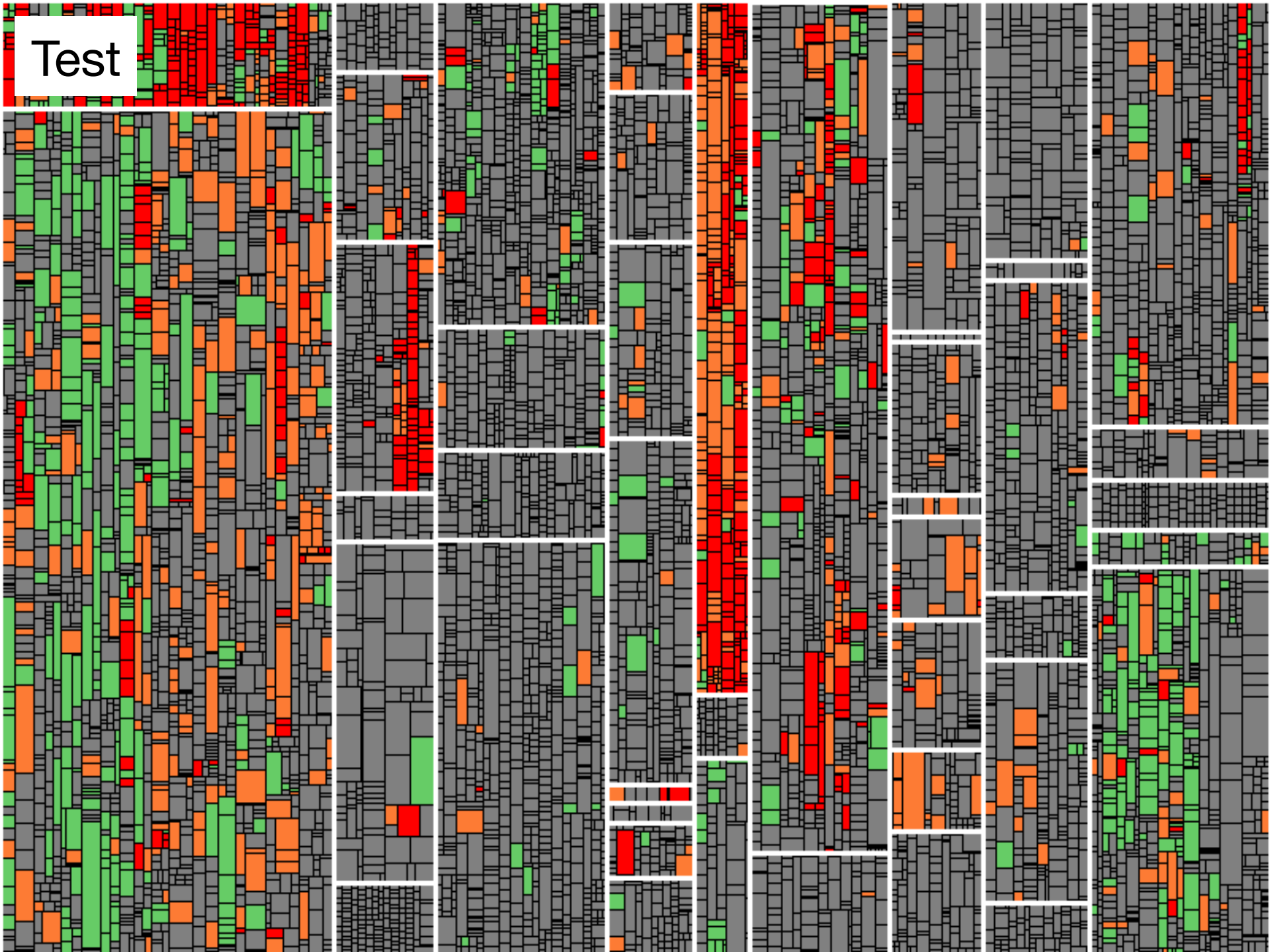
UAT



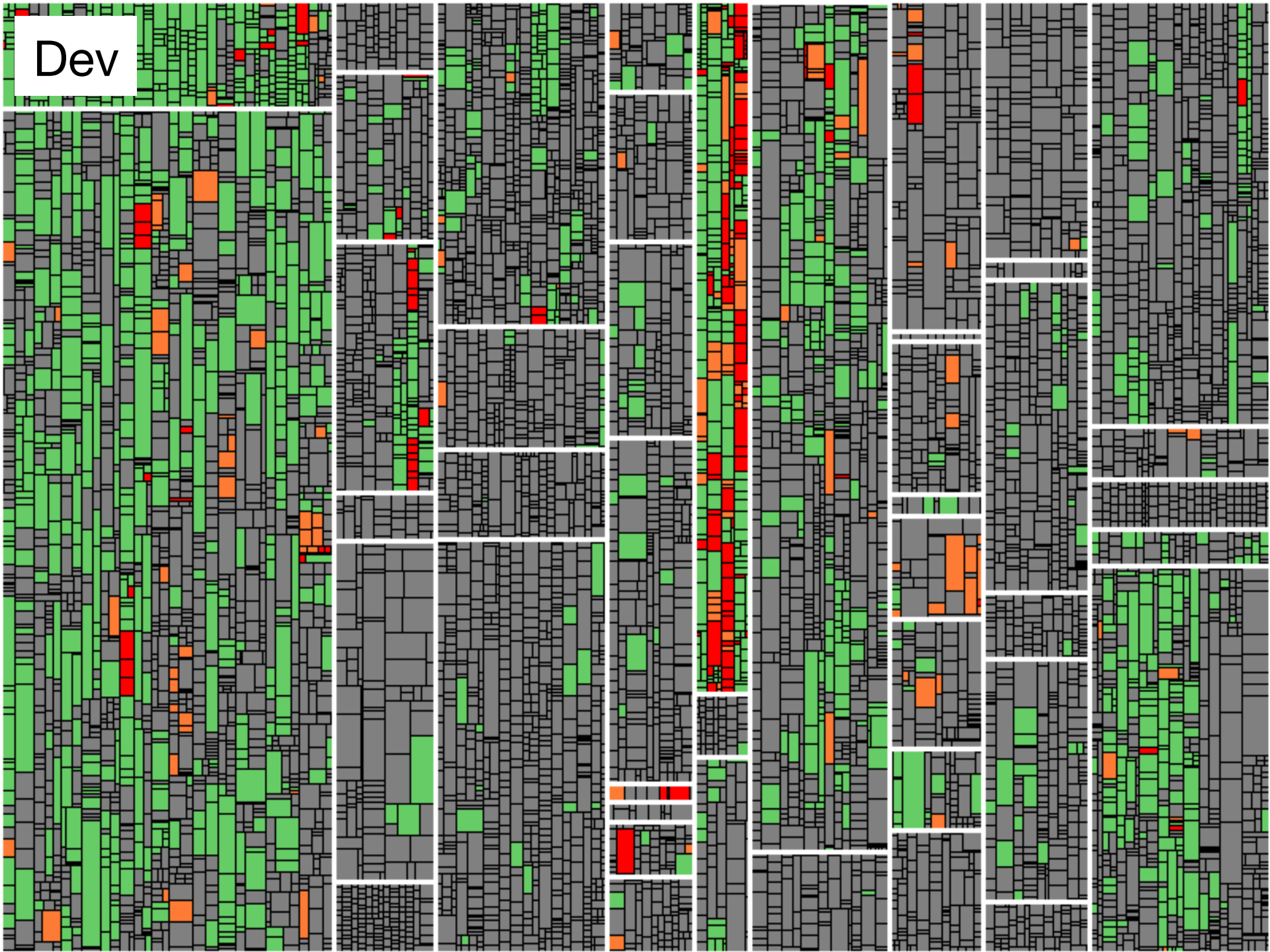
All



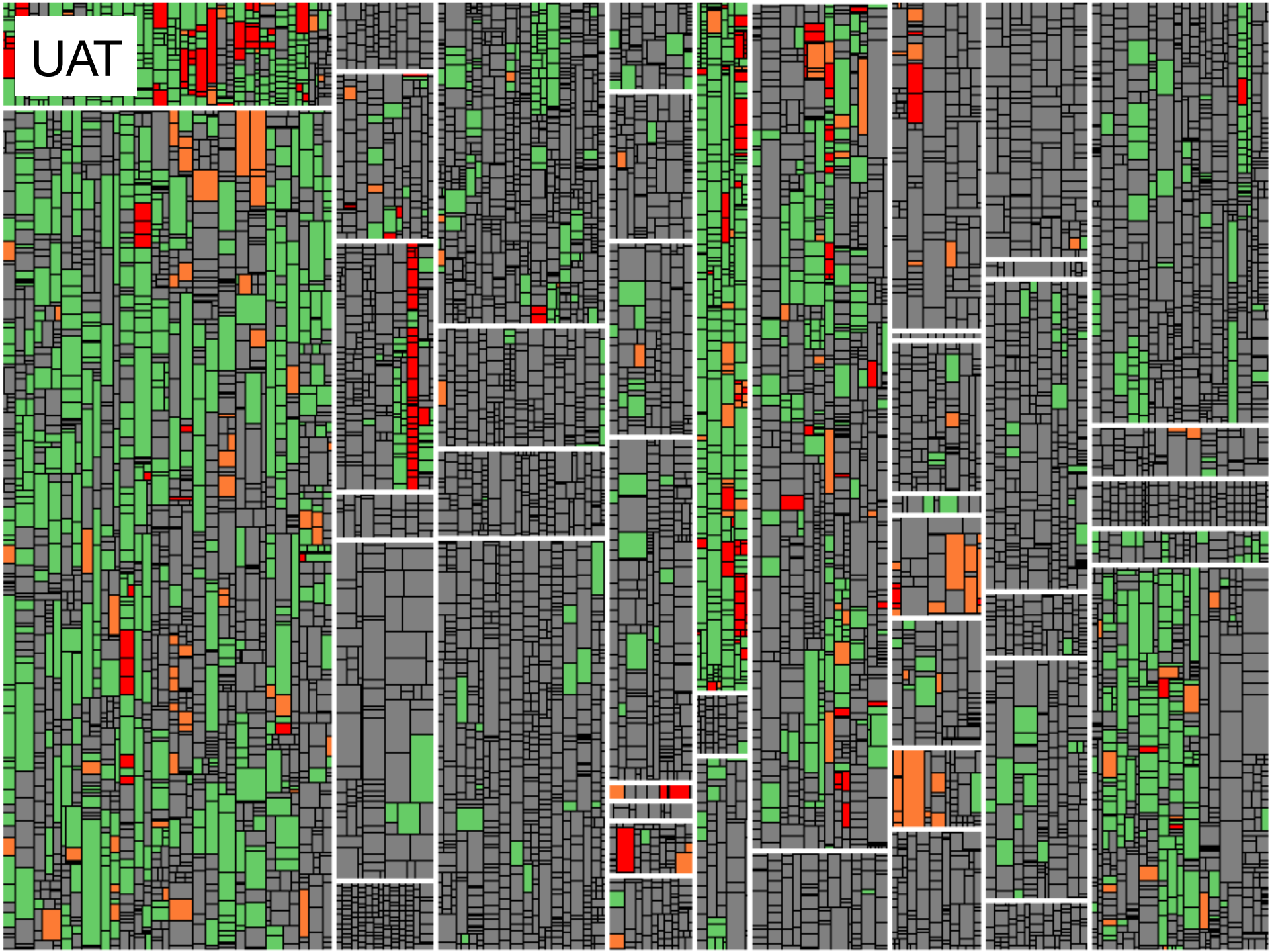
Test



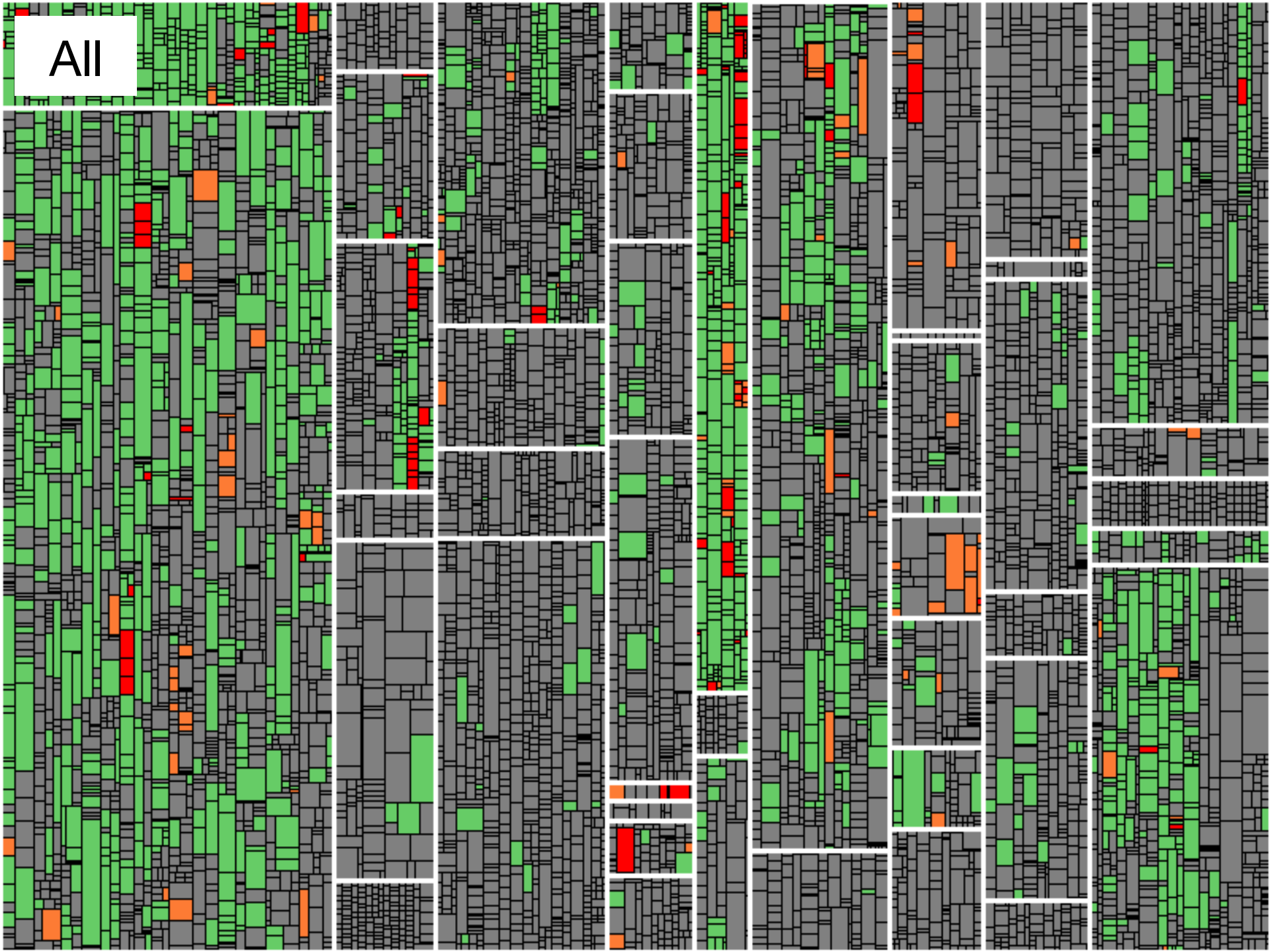
Dev



UAT

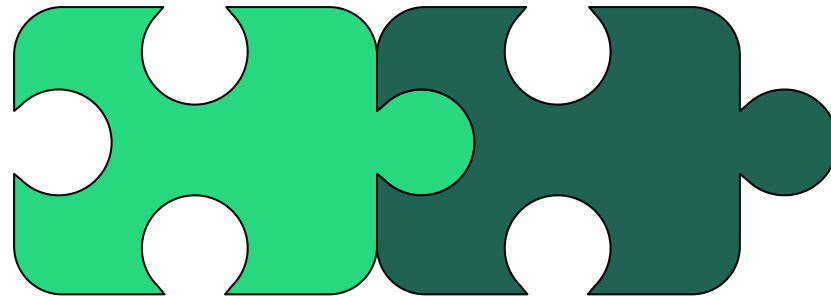


All

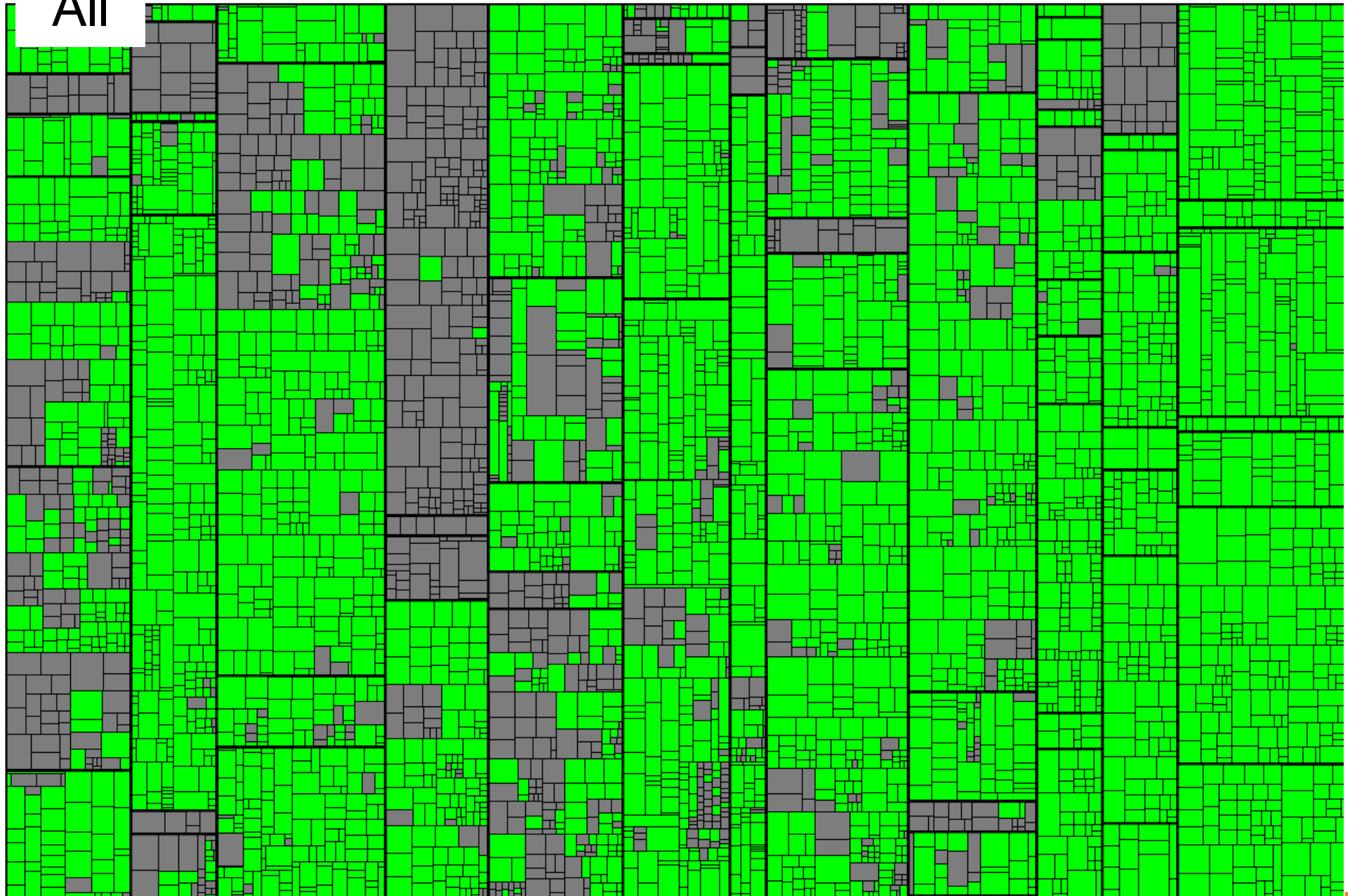


Regressions-Test: Testbreite optimieren

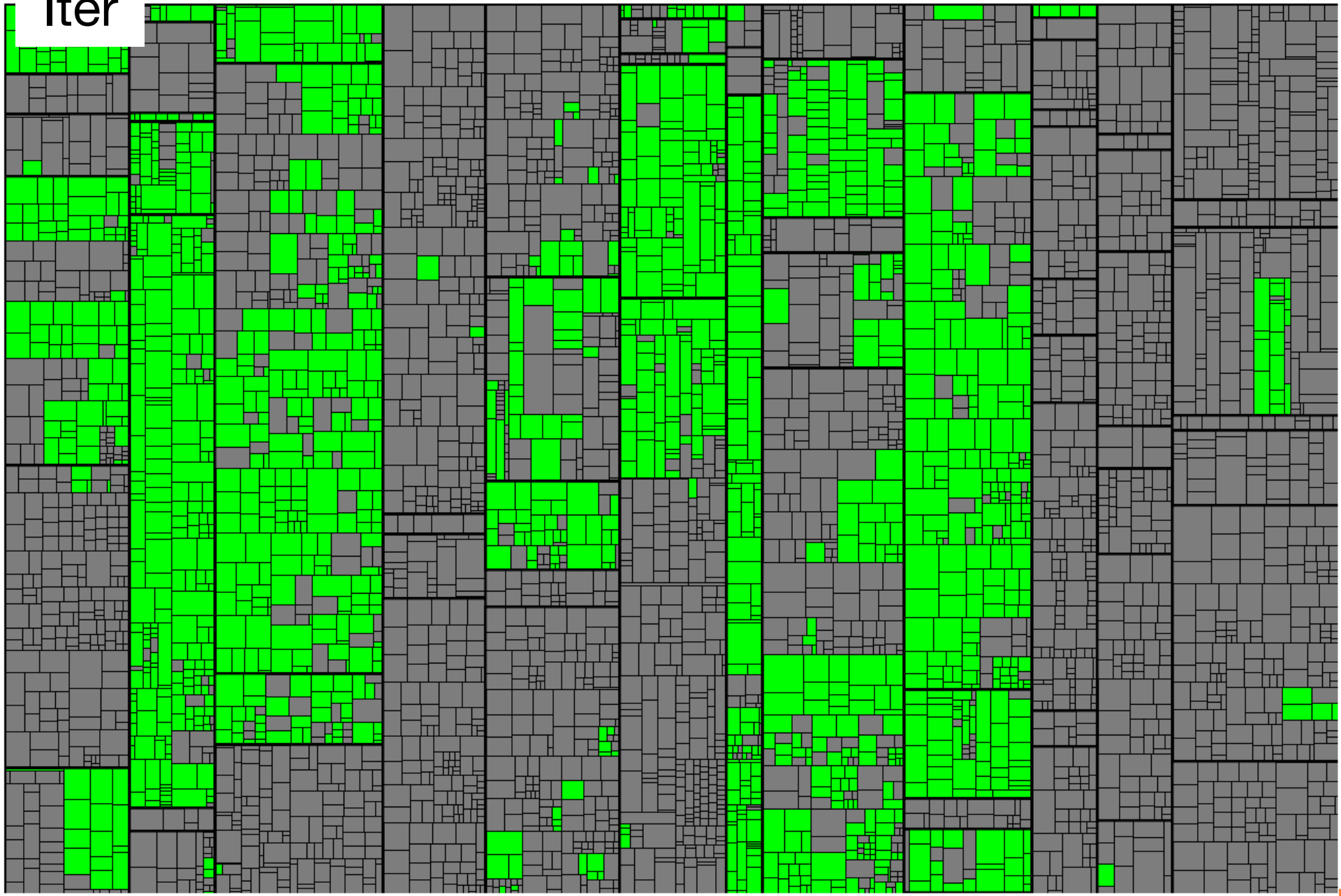
- Wie ergänzen sich verschiedene Teststufen wie Iterationstests, Unit-Tests, Regressionstests?
- Kann das Regressionstest-Set optimiert werden?



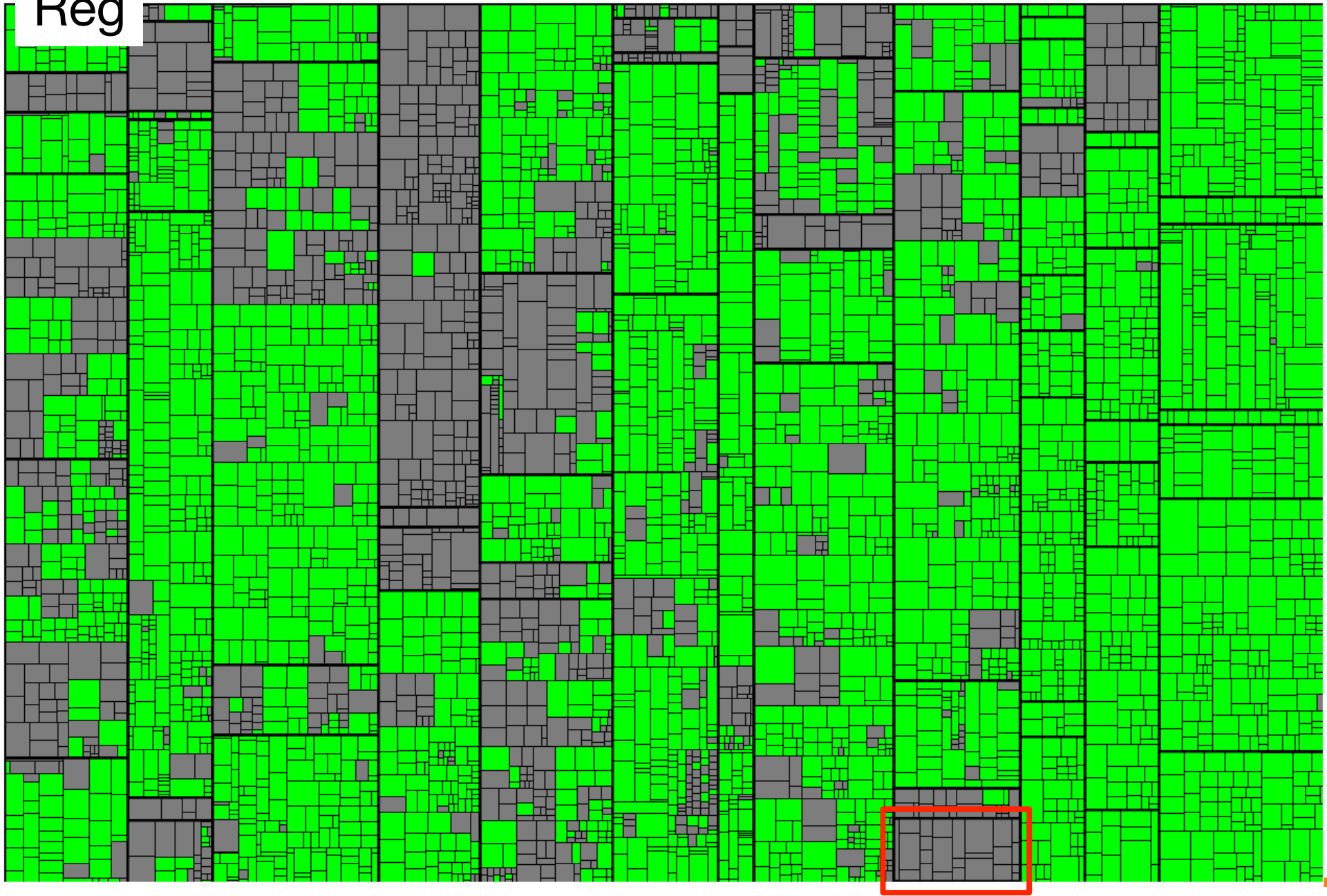
All



Iter



Reg

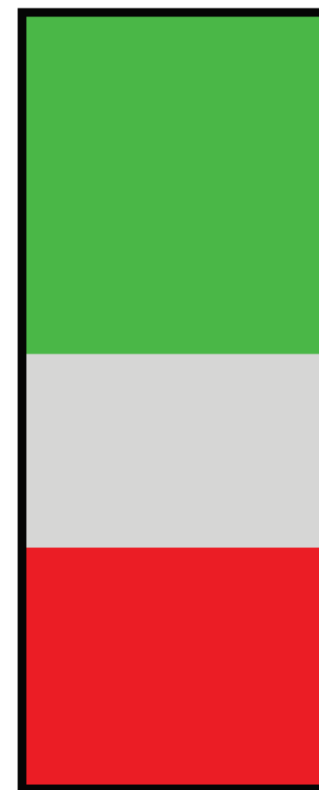
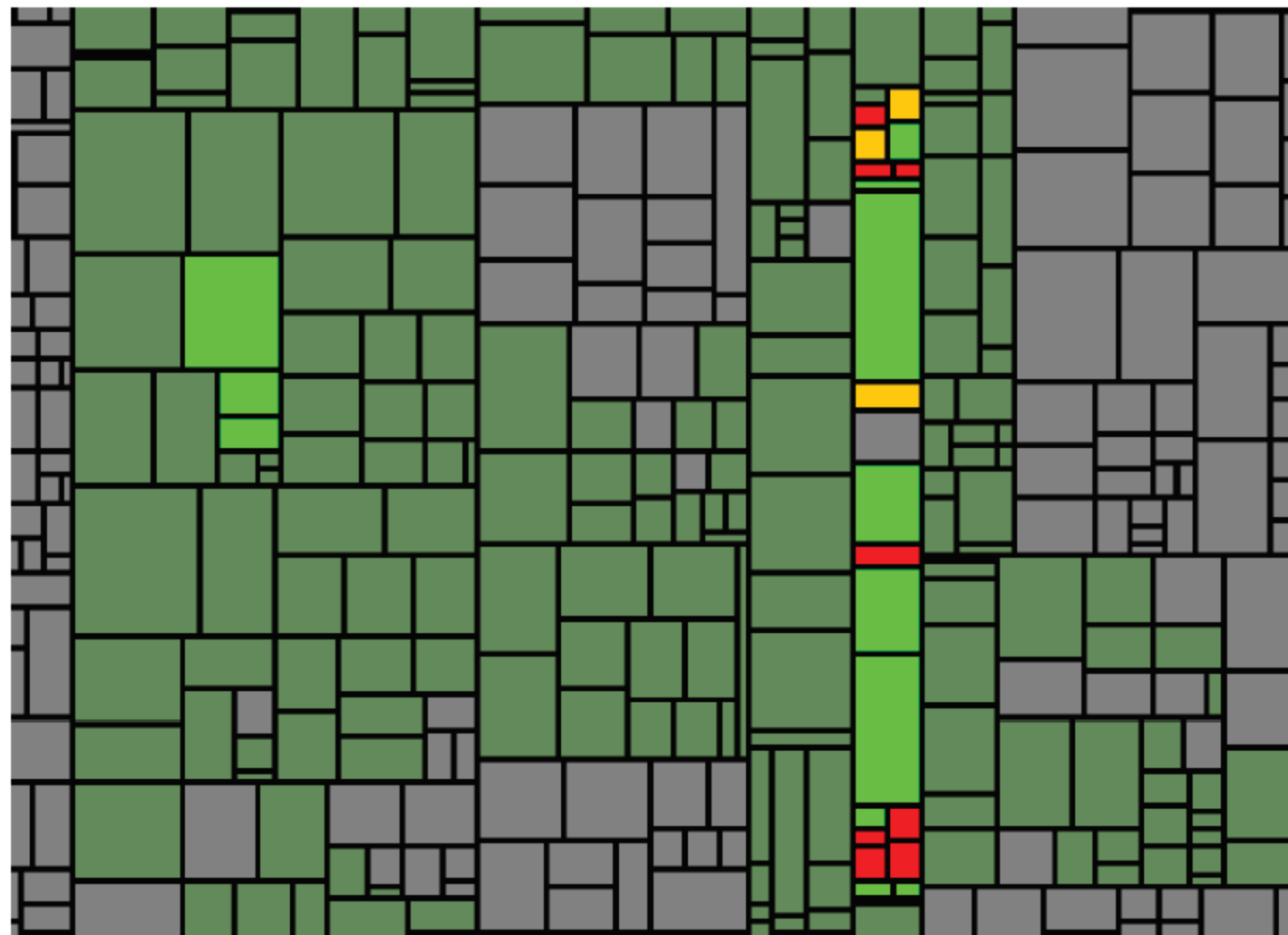


Agenda

- Motivation
- Test-Gap-Analyse
- Demo
- Erfahrungen aus der Praxis
- **Ausblick**

Change Request Coverage

Revision	Acti...	Author	Date	Message
24178		veihelm...	Montag, 25. April 2016 13:55:07	CR#7707: GREEN-RATING Closure compiler warning fixes
24177		veihelm...	Montag, 25. April 2016 13:42:41	CR#7707: GREEN-RATING Closure comment fix
24176		kinnen	Montag, 25. April 2016 13:30:14	CR#9702: GREEN
24175		kinnen	Montag, 25. April 2016 13:28:40	CR#9593: GREEN
24170		streitel	Montag, 25. April 2016 13:02:55	CR#9684: YELLOW fixing compiler warning
24169		streitel	Montag, 25. April 2016 11:43:53	CR#9684: YELLOW
24168		streitel	Montag, 25. April 2016 11:37:10	CR#9684: GREEN with Andi
24167		pawelka	Montag, 25. April 2016 11:28:47	CR#7707: fixing checkType warnings
24166		kinnen	Montag, 25. April 2016 09:15:58	Merged 2.0.x -> trunk
24165		haas	Freitag, 22. April 2016 18:27:50	CR#9697: YELLOW: Fix UI test (by using a new test case). Fix issue with Assume nodes in GFCs (they have no tok...
24163		baumeis...	Freitag, 22. April 2016 17:32:30	CR#9399: RED, incorporated review comment.
24162		goeb	Freitag, 22. April 2016 17:21:26	CR#9622: adjust handling of commit timestamps
24159		baumeis...	Freitag, 22. April 2016 17:08:18	CR#9399: GREEN, incorporated review comment.
24158		pawelka	Freitag, 22. April 2016 17:06:18	CR#0: Fixing architecture assessing if [ENTF] is clicked even if no element was deleted.
24157		pawelka	Freitag, 22. April 2016 17:05:01	CR#7707: fixing type check warning
24152		chasapis	Freitag, 22. April 2016 16:35:39	CR#9731: set minimum value for number input 0, because number of days cannot be negative.
24149		baumeis...	Freitag, 22. April 2016 15:49:19	CR#8487: YELLOW, incorporated review comments.
24148		haas	Freitag, 22. April 2016 15:40:30	CR#9697: RED: Temporarily disable refactoring suggestions in UI tests
24147		haas	Freitag, 22. April 2016 14:29:05	CR#9697: YELLOW: Improve data flow analysis for refactoring suggestions: * FOR, FOR EACH and WHILE loops (...
24146		haas	Freitag, 22. April 2016 14:22:48	CR#9697: YELLOW: Only consider candidates that have statements (i.e. not only META shallow entities, which occ...
24144		ashutoshb	Freitag, 22. April 2016 12:51:47	CR#9090: fixed the GUI-test.
24143		chasapis	Freitag, 22. April 2016 12:01:28	CR#8022: fixed nesting depth issue
24142		baumeis...	Freitag, 22. April 2016 11:57:54	CR#8487: YELLOW, incorporated review comments.
24141		baumeis...	Freitag, 22. April 2016 11:53:47	CR#8487: YELLOW, incorporated review comments.
24140		chasapis	Freitag, 22. April 2016 11:44:35	CR#8022: added functionality to all widgets that use a SingleMetricParameter, so that when the user edits the par...
24138		baumeis...	Freitag, 22. April 2016 11:37:24	CR#8487: YELLOW, renamed metric rules editor.
24137		baumeis...	Freitag, 22. April 2016 11:36:31	CR#8609: YELLOW, fix build.
24136		baumeis...	Freitag, 22. April 2016 11:18:17	CR#8609: YELLOW, added new notification utils method.
24135		gibtner	Freitag, 22. April 2016 11:03:51	CR#9364: changed status of files to YELLOW
24134		gibtner	Freitag, 22. April 2016 10:49:59	CR#9364: CommitTrackingService: - corrected comment - refactored: extracted methods
24133		baumeis...	Freitag, 22. April 2016 09:55:00	CR#8609: YELLOW, renamed tempalte parameters.

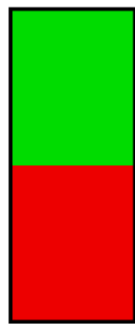


CR#7409

#c/a methods	16 [100%]
#not tested	5 [31,25%]
#exec in base	4 [25,00%]
#exec in test	7 [43,75%]



CR#6934



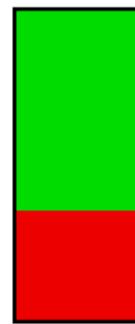
CR#7112



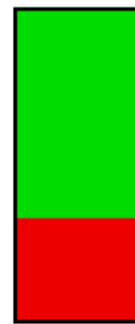
CR#6773



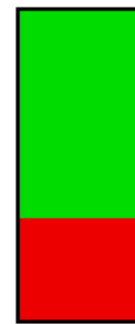
CR#8028



CR#8737



CR#6708



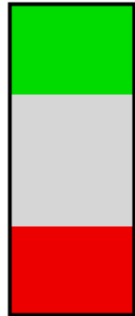
CR#8696



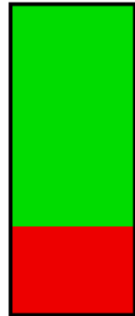
CR#7409



CR#8090



CR#8171



CR#6719



CR#8030



CR#8719



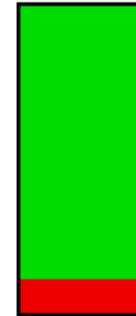
CR#7473



CR#7182



CR#7026



CR#8220



CR#7328



CR#7713



CR#8043



CR#8630



CR#7286



CR#6946



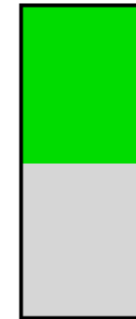
CR#6141



CR#7268



CR#7801



CR#7012



CR#8816



CR#8202



CR#8638



CR#8926



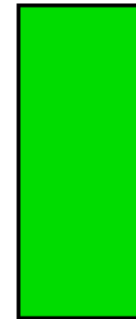
CR#6617



CR#6689



CR#6643

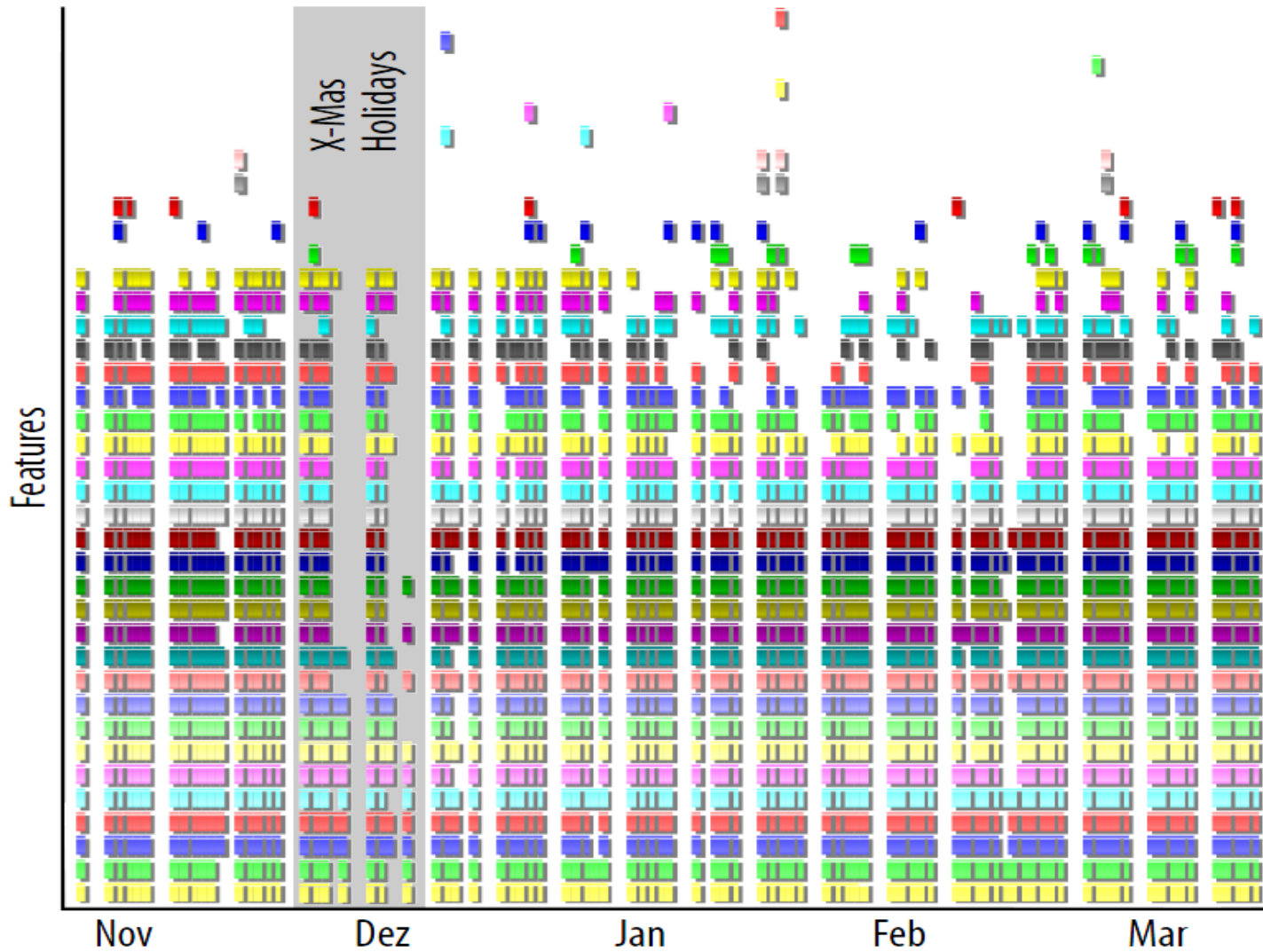


CR#6981



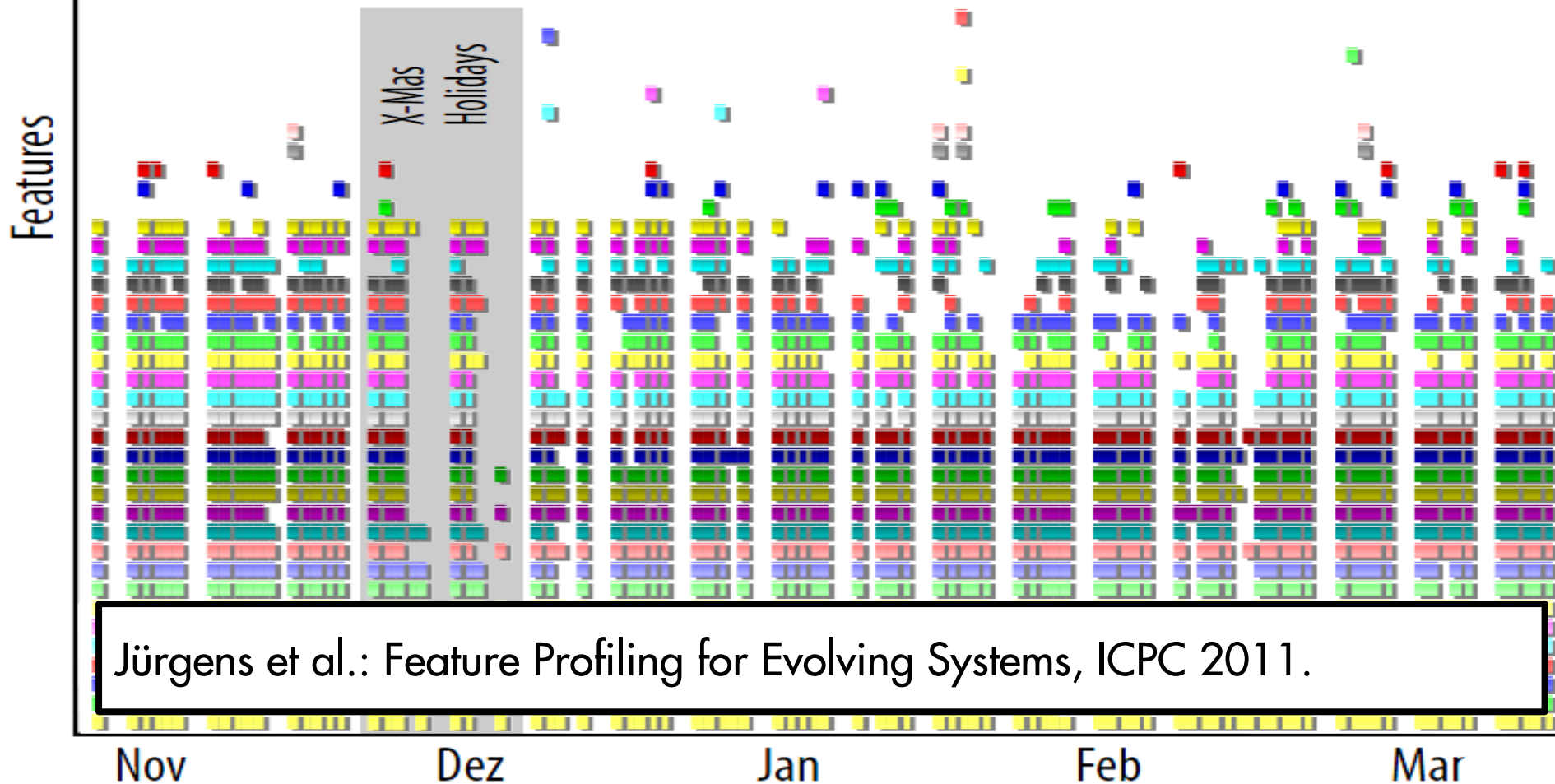
CR#7619

Nutzungsanalyse



28% der Features nicht benutzt (15/53)

- 5 völlig unerwartet,
- weitere 6 unerwartet für min. 1 Stakeholder



Fazit

- Code-Änderungen sind guter Fehlerprädiktor
- Test-Gap-Analyse macht Änderungen und Testaktivitäten transparenter und ermöglicht es **ungewollt** ungetestete Änderungen zu vermeiden.
- Wir unterstützen gerne bei der Evaluierung.

Kontakt

Ich freue mich auf Diskussionen!

Für eine Evaluierungslizenz von Teamscale (www.teamscale.com)
bitte bei mir melden.

Bei Interesse an einem Job in diesem Bereich, bitte auch melden 😊

Dr. Dennis Pagano · pagano@cqse.eu · +49 1590 4062957

🐦 @dennispagano

CQSE GmbH

Lichtenbergstraße 8

85748 Garching bei München

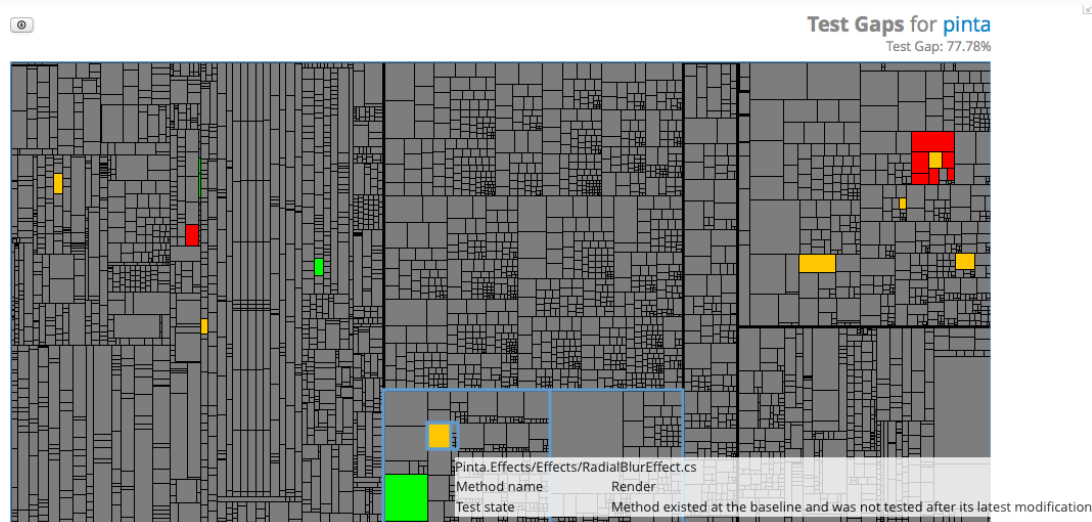
Software Quality Blog

Bridge your Test Gaps with Teamscale

Posted on 04/27/2016 by Dr. Dennis Pagano

Many companies employ sophisticated testing processes, but still bugs find their way into production. Often they hide among the subset of changes that were not tested. Actually, we found that untested changes are **five times more error prone** than the rest of the system.

To avoid that untested changes go into production, we came up with *Test Gap analysis*—an automated way of identifying changed code that was never executed before a release. We call these areas *test gaps*. Knowing them allows test managers to make a conscious decision, which test gaps have to be tested before shipping and which ones are of low risk and can be left untested.



Software Quality Blog

Take control of your testing process!

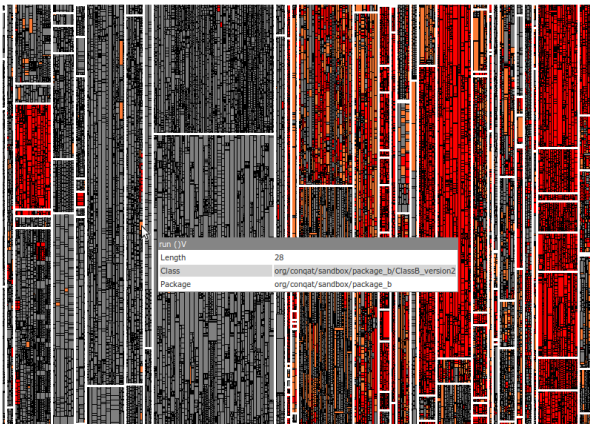
Posted on 10/08/2014 by Fabian Streitl

Testing is an integral part of a software product's life-cycle. Some people prefer executing their tests continuously while they develop, others have a separate testing phase before each release. The goal, however, is always the same: finding bugs. The more, the better.

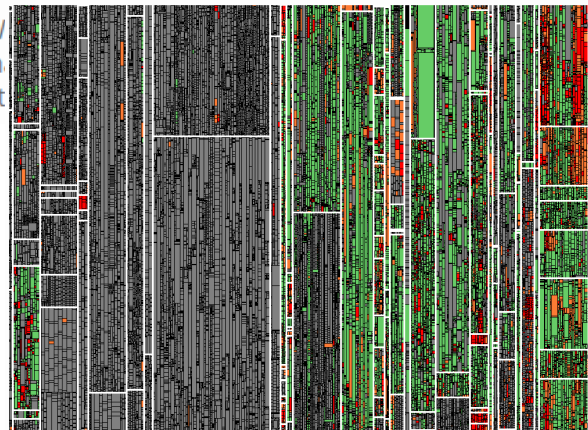
Unfortunately, the time available for testing and for writing new tests is limited. At some point, you have to get on with development, ship your software and be confident it contains no serious flaws. Often, more tests exist than can be run in the available time span. This is especially true if your tests are executed manually, as is the case for many of our customers. Then the question becomes: which tests do I select to find the most bugs before the release?



Research h
errors thar
true wheth



likely ("Did W
ire that all ch
hole test suit



this is

Testing Changes in SAP BW Applications

Posted on 04/29/2015 by Dr. Andreas Göb

As my colleague Fabian explained a few weeks ago, a combination of change detection and execution logging can substantially increase transparency regarding which recent changes of a software system have actually been covered by the testing process. I will not repeat all the details of the Test Gap Analysis approach here, but instead just summarize the core idea: Untested new or changed (informative) elements. Therefore it makes sense to use

Several times
In the main
from Python
containing
may provide
Analysis in

