



ورقة بحثية استكشافية

التاريخ التحريري: 22 أكتوبر 2026

من التقرير إلى العقد التنفيذي: دراسة مفاهيمية وتجربة حسابية في قابلية تصدير معرفة الأعطال البرمجية

From Incident Reports to Executable Contracts: A Conceptual and Computational Study of Exportable Software-Failure Knowledge

التصنيف العلمي: ورقة بحثية مفاهيمية مدعومة بمراجعة أدلة أولية محددة النطاق، وبمُبرهن تطبيقي برمجي منقذ على حالات وبيئات اصطناعية. **الحالة:** أنجز استخراج المصادر المحددة، وتصميم النموذج، وكتابة الشيفرة وتشغيلها وفحص ملفات نتائجها. **لم يُنجز:** استخراج آلي موثوق من تقارير حقيقية باستخدام نموذج لغوي، أو تجريب على منتجات عملاء، أو دراسة اقتصادية فعلية لشركة سعودية، أو اختبار تعمية مستقل، أو تحكيم علمي.

الملخص العربي

تبحث الدراسة إمكان تحويل الخبرة الموثقة في حوادث البرمجيات والأنظمة الموزعة إلى أصل تقني قابل للتشغيل والنقل والتسويق خارج بلد تطويره. لا تنطلق من فرضية أن تعرض دولة ما لضغط إقليمي ينتج تلقائيًا صناعة قابلة للتصدير؛ بل تفصل بين أربعة مستويات: وصف واقعة، واستخراج آلية فشل يمكن تعميمها بصورة مقيدة، وكتابة اختبار تنفيذي ذي قاعدة حكم قابلة للتدقيق، وإثبات أن المشتريين الأجانب يحصلون منه على قيمة تفوق تكلفة التركيب والصيانة. تربط المراجعة ثلاث وقائع موثقة: تعطل «كلاودفلير» (Cloudflare) في 18 نوفمبر 2025 بسبب تضاعف ملف خصائص بعد تغيير صلاحيات، وتعطل خدمات «غوغل كلاود» (Google Cloud) في 12 يونيو 2025 بسبب بيانات حصص ذات حقول فارغة وتكرار استعادة مُرهق، وانقطاع شبكة «إيه تي أند تي» (AT&T Mobility) في 22 فبراير 2024 بعد خطأ تهيئة شبكية واسع الأثر. ثم تنفذ دراسة اصطناعية على 16 عقد اختبار و4 تطبيقات صُنعت لهذا الغرض، بإجمالي 64 نتيجة، فتجد 7 و12 و11 و16 موافقة على التوالي لنتائج الاختبارات

المرجعية. هذه الأرقام ناتجة بحكم تصميم البرمجيات وعقودها، ولا تقيس فعالية في عالم الإنتاج أو احتمال منع الأعطال. تخلص الدراسة إلى أن الفرضية التجارية الجديرة باختبار مستقل ليست «استخدام الذكاء الاصطناعي لصنع اختبارات فوضى»، فهذه ممارسة موجودة، بل قياس قابلية نقل الدلالة وقاعدة الحكم بين بيئات مختلفة، وتكلفة نقلها، وتجدد طلب المشتريين عليها.

الكلمات المفتاحية: العقود التنفيذية للفشل (Executable Failure Contracts)، انتقال الدلالة (Semantic Transfer)، اختبار الخصائص (Property-Based Testing)، التجربة المقيّدة (Bounded Experiment)، برمجيات الاستدلال (Inference Software)، هندسة الفوضى (Chaos Engineering)، اقتصاديات وحدة المنتج (Product Unit Economics)، الخدمات البرمجية للتصدير (Software Services Exports).

English abstract

This conceptual and computational working paper examines whether documented software failures can be transformed into portable, independently inspectable test artifacts with exportable commercial value. It distinguishes incident evidence, failure-mechanism abstraction, executable-oracle construction, cross-system semantic transfer, and demonstrated foreign willingness to pay. Three primary records anchor the inquiry: Cloudflare's November 2025 feature-file propagation outage, Google Cloud's June 2025 quota-policy failure and recovery congestion, and the February 2024 US AT&T mobile-network outage investigated by the FCC. A reproducible, fully synthetic Python demonstration manually specifies 16 toy test contracts spanning input schemas, bounded feature files, authenticated versioning and retry limits. Four intentionally constructed toy adapters yield 64 executions and respectively match 7, 12, 11 and 16 declared oracle outputs. The exercise demonstrates contract encoding and repeatability, not predictive incident prevention or product-market fit; all outcomes are consequences of artificial design choices. The paper proposes prospective tests for cross-platform semantic transfer, oracle accuracy, false positives, incremental value over established chaos-testing products, integration cost and actual foreign renewals. The central claim is a falsifiable opportunity hypothesis, not evidence that a Saudi deep-tech resilience-export industry already exists.

Keywords: executable incident contracts; semantic transfer; chaos engineering; test oracle; fault injection; Saudi technology exports; reproducibility; product-market fit

1. المسألة العلمية وما الذي يتغير لو عرفنا جوابها؟

تدعو أدبيات موثوقة الأنظمة إلى فهم الأعطال والاستعداد لها، لكنها لا تثبت وجود سوق مستقلة لأي أداة جديدة توصف بأنها تزيد المرونة. و«صناعة تصديرية» هنا مفهوم اقتصادي أشد تحديداً من مهارة فريق هندسي محلي في حل مشكلاته. نريد اختبار الفرضية الآتية:

توجد فئات من معرفة الإخفاقات البرمجية يمكن استخراجها من حوادث موثقة أو بيانات مصرّح بها، وتمثيلها كعقود اختبار مستقلة نسبياً عن المورد، ونقلها بين نظم عملاء مختلفين بتكلفة حدية معقولة، ثم بيعها بصورة متكررة إلى مشترين من خارج السعودية يدفعون بسبب منفعة هندسية مثبتة.

هذه فرضية مركبة، ولا يكفي إثبات جزء منها لإثبات مجموعها. فقد تكون هناك قابلية تقنية لصنع اختبار ولكن من دون انتقال دلالي بين المنصات؛ وقد تكون هناك قابلية انتقال مع تخصيص باهظ لكل مشروع؛ وقد يتمكن فريق من بيع خدمات خبراء دولياً دون أن يتحقق نموذج برمجيات منتج متكرر. والإجابة مهمة لأنها تغير قرار التخصص والاستثمار: هل يجدر تمويل مختبر لبناء محوّلات ومواصفات قابلة لإعادة الاستخدام، أم الاستثمار في فرق استشارات متخصصة، أم الاكتفاء باستخدام الأدوات القائمة؟

وحدة التحليل العلمية: عقد اختبار محدد، ينفذ على نظام مُعرّف، تحت مجموعة شروط، ويصدر نتيجة مقارنة بقاّض معلن. **وحدة التحليل التجارية:** تعاقد عميل أجنبي مستقل على منتج أو تجديده خلال دورة بيع وصيانة محددة. لا نساوي بين الوحدتين.

2. ما الموجود فعلاً قبل المشروع المقترح؟

2.1 تجارب الأعطال أدوات قائمة وليست اكتشافاً سعودياً مفترضاً

تشرح مبادئ هندسة الفوضى (Principles of Chaos Engineering) الاختبار حول فرضية لسلوك مستقر، والتجريب المبني على المشاهدة، وتقليل نطاق أثر التجربة (Blast Radius). وهي لا تقول إن التجربة تُصنع من تقارير عامة بلا معرفة بالنظام. وتتيح خدمة «أمازون لإدخال الأعطال» (AWS Fault Injection Service; AWS FIS) مكتبة سيناريوهات منذ 2023، كما يصف دليل نشرته الشركة في S06 2024. ولها شروط إيقاف مرتبطة بإنذارات محددة، وليست مجرد مفتاح لإيقاف الآلات S08. وفي 2025 نشرت «أمازون» (Amazon) مثلاً هندسياً يستخدم نماذج لغوية (Large Language Models; LLMs) لتوليد قوالب تجارب من أوصاف طبيعية على بيئتها S07. كل ادعاء بأن مجرد «تحويل النص إلى قالب اختبار» جديد أو خالٍ من المنافسة يناقض هذه الأدلة.

كذلك يعرض مشروع «جيبسن» (Jepsen) منذ 2013 تقارير عن فحص الاتساق في قواعد بيانات وصفوف ورسائل موزعة. يستعمل تطبيقات فعلية وتوليد عمليات وأعطالاً مثل تقسيم الشبكة وانحراف التوقيت، ثم يفحص تاريخ المخرجات في ضوء نماذج اتساق معلنة S09. ولذلك ينبغي تحديد موضع الإضافة الممكنة بدقة: قد يكون المنتج

عرضًا موحدًا لمكتبة حالات متغيرة بتحديثات الحوادث، أو تحققًا دلاليًا عبر موردين، أو إثباتًا قابلاً لإعادة التشغيل للمشتري. لكل بديل منافسون مختلفون، ولا نستطيع تقرير فجوة السوق قبل مراجعة مشترياتهم واختبار المقارنات.

2.2 ثلاث عائلات من آليات الفشل في الأدلة الأولية

يعتمد التأطير على السجلات التالية، وهي لا تشكل عينة ممثلة لتكرار كل الأعطال التقنية:

الحادثة	ما وثّقه المصدر	آلية مجرّدة مقترحة	ما لا نعرفه
18 نوفمبر 2025، «كلاودفلير» (Cloudflare)	ملف خصائص تضاعف بسبب اختلاف نتائج استعلام بعد تغيير صلاحيات؛ تجاوز الحد الذي يقبله قارئ الملف، وتذبذبت الخدمة <u>S01</u>	توسع غير متوقع في مخرجات وسيطة موزعة، مع قارئ ذي افتراض حجم مخفي	هل يكفي اختبار إزالة التكرار في نظم أخرى؟ لا؛ قد يختلف المعنى وترتيب العناصر
12 يونيو 2025، «غوغل كلاود» (Google Cloud)	بيانات حصص منقوصة أحدثت استثناء برمجياً في مسار موزع عالمياً؛ ضغطت إعادة المحاولة على مخزن خلفي أثناء التعافي <u>S02</u>	تنشيط مسار غير ممتحن مع بيانات ناقصة ثم ازدحام الاستعادة	لا تتوافر لدينا نسخة تنفيذية من النظام الداخلي لتكرار العطب الأصلي
22 فبراير 2024، «إيه تي أند تي» (AT&T Mobility)	خطأ إعداد عنصر شبكي عطل الاتصال على نطاق وطني؛ 92 مليون مكالمة حجبت وأكثر من 25 ألف اتصال طوارئ لم يصل حسب تقرير الجهة التنظيمية <u>S03</u>	انتشار أثر تغيير عنصر تحكم وإيقاع إعادة ارتباط واسع	لا تُحدد هذه المعطيات ماهية تجربة أمانة كاملة على مشغّل مختلف

المصدر الأول والثاني تقريران يكتبهما المشغّلون عن أنفسهم. للناشر معرفة تقنية مباشرة لكنه طرف ذو مصلحة في رواية الحادثة؛ لم نقارنها بسجلات تشغيل خام غير منشورة. المصدر الثالث تحقيق جهة تنظيمية مستقل نسبياً يستند في جزء من أرقامه إلى ما أفصح عنه المشغّل. تبقى هذه الحدود في سجل الادعاءات.

2.3 دليل سياقي على وجود طريق للتصدير، لا سوق لهذا المنتج

ذكرت وكالة الأنباء السعودية (Saudi Press Agency; SPA) مشاركة 28 شركة تقنية واتصالات وطنية ضمن جناح «صناعة سعودية» في معرض «جيتكس أفريقيا» (GITEX AFRICA) في أبريل 2025 S04، كما نشرت خبر توقيع مذكرة تفاهم بين «علم» (Elm) و«هواوي» (Huawei) خلال معرض «ليب إيست» (LEAP East) في يوليو 2026 S05. هذه وثائق عن الحضور المؤسسي والأنشطة الدولية فقط، ولا تتضمن قياساً مستقلاً للصادرات المحصلة في قطاع الاختبارات المقترح، أو اشتراكات متكررة، أو قاعدة عملاء أجنبية، أو هامش ربح هذا النوع من المنتجات. تعد مؤشراً إلى إمكانية البحث عن قنوات مبيعات، لا متغيراً تابعاً يثبت نجاح الأطروحة.

3. تعريف الشيء الذي يفترض تصديره

نعرّف عقد الفشل التنفيذي (Executable Failure Contract) المقترح بوصفه سجلاً ذا خمسة عناصر ضرورية:

1. المنشأ (Provenance). رابط السجل الذي استُخرجت منه الفرضية، وموضع الفقرة، ودرجة استقلال المصدر، وما إذا كانت العلاقة سببية معلنة أو تفسيرًا استنتاجيًا.
2. الشروط المسبقة (Preconditions). خصائص التطبيق والمخزن والبروتوكول التي تجعل التجربة ذات معنى؛ وقيودًا تمنع إسقاطها على نظام مختلف بصورة مضللة.
3. المنبه التجريبي (Stimulus). إدخال اصطناعي، أو تغيير مصرح به في بيئة معزولة، أو نمط عمليات مصطنع مع نطاق محدد ومدة وحدود توقف.
4. قاضي السلوك (Test Oracle). خاصية ملاحظة أو ثابت وظيفي يفصل نتائج مقبولة من نتائج مرفوضة، مع تحديد كيفية ملاحظة السلوك وعدم اليقين.
5. حزمة الشاهد (Evidence Package). نسخ البرمجيات، إعدادات التشغيل، الأذونات، سجل الأحداث والاختبار، النتائج، المفقودات، وتفسير ما لا يجوز استنتاجه.

مهم أن تكون الوحدة قابلة للنقل من خلال تمثيل وسيط (Intermediate Representation; IR)، لكن النقل ليس مساواة آلية بين أسماء الحقول. إذا افترض عقد اختبار أن حذف تكرار عنصرين لا يغير معنى الملف، فإن تحويله إلى خدمة يُعد فيها التكرار مقصودًا يفسد العقد. ولذلك نحتاج وصفًا وظيفيًا وشروطًا للدلالة لا مجرد تحويل صيغ (Serialization).

نقترح هيئة توضيحية، لا معيارًا معتمدًا:

```
incident_spec_version: proposed-v0
provenance:
  source_id: S01
  claim_boundary: documented mechanism; not original payload
applicability:
  data_semantics: order-insensitive feature set
  environment: isolated local simulator
stimulus:
  type: repeated_features
  records: [a, b, a, b, a, b, a]
oracle:
  acceptable: ACCEPT after semantic deduplication
  forbidden: uncaught parser exception
safe_execution:
  production_connection: false
  stop_after_seconds: 10
  rollback: discard isolated fixture
limits:
  does_not_reproduce: Cloudflare production code or actual outage
```

وجود اسم «كلاودفلير» في حقل المنشأ لا يعني أن العينة تسوّي الاختلافات التقنية بينها وبين أي نظام آخر. وإذا لم يثبت المستعمل أن الملف المستهدف مجموعة لا قائمة ذات ترتيب دلالي، يجب رفض تطبيق هذا العقد عليه؛ الامتناع عن الترجمة (Abstention) أفضل من إصدار نتيجة واثقة بلا معنى.

4. المشكلة الأصعب: هل ينتقل المعنى أم تنتقل الأوامر فقط؟

تقوم قابلية النقل (Portability) على ثلاثة مستويات متباينة:

- **نقل الصيغة (Syntactic Portability):** يمكن قراءة الملف وتحويله إلى قالب صالح لغويًا في محرك آخر.
- **نقل العملية (Operational Portability):** يمكن تنفيذ الاختبار داخل حدود سلطة واضحة مع تسجيل البيانات المطلوبة دون تعريض نظام حي للخطر.
- **نقل الدلالة (Semantic Portability):** تختبر العملية الجديدة الخاصية نفسها، وتتعامل مع الحالات المقبولة والمرفوضة على نحو منسجم مع وظيفة التطبيق الجديد.

قد تنجح المرحلتان الأوليان وتفشل الثالثة. مثال: محاكاة انقطاع منطقة سحابية قد تفضي إلى انقطاع شبكة الاختبار، لكنها لا تماثل تدهور سياسة نسخ موحدة تتوزع إلى مناطق متعددة. كذلك فإن سجل تتبع (Trace) غنيًا بالأزمنة لا يكشف وحده فقدان معاملة أو تكرارها، ما لم نحدد علاقةً منطقية بين العمليات المتزامنة.

لضبط السؤال، ليكن لدينا نظامان تجريبيان A و B ، وأثران تنفيذيان τ_A و τ_B ، وتحويلان مجردان للمشاهدات α_A و α_B ، وخاصية مشتركة φ . الادعاء بالنقل الدلالي يتطلب تقريبًا أن نقارن حكم الخاصية على تمثيلات آثار متكافئة ضمن شروط التطبيق، لا أن نلاحظ أن الاختبارين انتهيا دون استثناء. يمكن كتابة مطلب المعنى بصورة رمزية:

$$\text{MeaningMatch}(A, B, \varphi) := \text{Comparable}(\alpha_A, \alpha_B, \varphi) \wedge \text{OracleSound}(A, \varphi) \wedge \text{OracleSound}(B, \varphi).$$

هذه صيغة مفاهيمية، لا برهانًا على تكافؤ برامج أو قاعدة قرار آلية. وتوجد خصائص يستحيل الحفاظ عليها خارج بيئة بعينها، مثل دلالات معاملات مالية أو أدوات قانونية تختلف في نظام آخر.

5. تصميم التجربة المنفذة فعليًا

5.1 الهدف المحدود

لم نُشئ مترجم حوادث لغويًا ولم ننفذ اختبارات على «كلاودفلير» أو «غوغل» أو «إيه تي أند تي»، ولم نستورد عينات من أنظمتها. بنينا مبرهنًا تطبيقيًا (Proof-of-Concept Demonstrator) بسؤال واحد: هل يمكن تدوين مجموعة صغيرة من ثوابت السلوك بطريقةٍ تنتج مخرجات قابلة للمقارنة بين تطبيقات مستقلة في بيئة محلية حتمية؟

كُتبت الأمثلة يدويًا. احتوت على ستة عشر عقدًا: أربع حالات لكل عائلة من أربع عائلات. جُعِلت كل حالة تحتوي `observed_property` و `case_id`, `family`, `description`, `stimulus`, `expected`. ثم كُتبت أربع وحدات تنفيذ اختبارية تحمل أربع مجموعات متفاوتة من الحواجز. البرنامج بلغته «بايثون» (Python) ويستعمل مكتبتها القياسية فقط؛ لا يتصل بالشبكة. ويشمل السجل الرقمي بصمة تجزئة (SHA-256) لتكوينات حالات الاختبار كي يُكتشف تبديل المدخلات دون الإعلان عن تحقق أمني خارجي.

المدخلات والسياسات (Policy Schema). حصة رقمية سليمة، وحصة مفقودة، وحصة سالبة، وحصة غير قابلة للتحويل إلى رقم. الشرط المقترح: إذا تعدّد تفسير البيانات يُرفض الإدخال بصورة متعمدة، لا يُترك استثناء غير معالج يقتل مسار المعالجة. ولا تعني هذه السياسة أن كل خدمة إنتاجية يجب أن تُغلق عند كل نقص حقوق.

ملف الخصائص (Feature Bundle). مجموعة صغيرة، وتكرار مسموح، وتكرار يبلغ حجمًا خامًا فوق السقف لكن بعد إزالة التكرار يصبح ضمن السقف، ومجموعة فريدة تفوق السقف. في المثال الاصطناعي فقط عرّفنا الدلالة بأن ترتيب العناصر وتكرارها لا يغيّر المعنى؛ لذلك يصح حذف التكرار. إذا لم تكن هذه الصفة مثبتة لنظام حقيقي فلا يجوز النقل.

النسخ والتوثيق (Update Version). تحديث جديد موثق، وإعادة إرسال نسخة مساوية موثقة، ونسخة قديمة موثقة، ونسخة جديدة لكن ذات توقيع غير صالح. الشرط المقترح ألا يطبّق البرنامج تحديثًا متراجعًا أو غير موثق، مع قبول الإعادة المطابقة حيث تكون عاقبة ذلك معروفة. لا تمثل البنية الحقيقية للتوقيعات أو لإدارة المفاتيح؛ الحقل `signature_ok` قيمة اصطناعية جاهزة.

إعادة المحاولة المحدودة (Bounded Retry). نجاح مباشر، وثلاثة إخفاقات قبل النجاح في المحاولة الرابعة، وإخفاق ممتد لاثنتي عشرة محاولة، وستة إخفاقات متوالية. الحدّ في المثال أربع محاولات إجمالية. تقارن قاعدة الحكم بين نجاح مشروع داخل الحد والتوقف عند استنفاده. لا تنفذ الشيفرة شبكة أو إرجاء عشوائيًا حقيقيًا، ولا تقيس تأثير الزحام على مخزن بيانات.

5.3 أربع تطبيقات اصطناعية

1. **بلا حواجز (Naive).** يحاول تحويل الحقول مباشرة، ويستخدم طول القائمة الخام وحدًا ذا استثناء، ويقبل نسخ التحديث، ويعيد المحاولة حتى عدد مبالغ فيه.
2. **حراسة المدخلات (Input Guard Only).** يضيف تحققًا من نوع الحصص وقيمتها ومن نسخ التحديث والتوقيع؛ ويبقى سلوك الملف وإعادة المحاولة هسًا.
3. **حراسة حدود الموارد (Capacity Guard Only).** يضيف معالجة تكرار الملف وحدّ إعادة المحاولة؛ ويبقى تحقق السياسة والنسخ هسًا.
4. **حواجز مجتمعة (Combined Guards).** يجمع التحقيقات الموضوعة صراحةً في عقود المثال.

كل تطبيق صُنِعَ عمدًا لإظهار اختلاف محدود. ليس الأول شركة حقيقية سيئة، ولا الرابع تصميمًا مثبتًا آمنًا. لأن بناء الاختبارات والحواجز تمّ في مشروع واحد، لا يصلح نجاح الرابع دليلًا إحصائيًا على تفوق منهجية تطوير بعينها في العالم. ويمكن تصور حالات جديدة ستفشلها جميعًا.

5.4 التنفيذ والنتائج المسجلة

شُغِلَت `analysis/reproduce.py` وأُنْتُجَت `contract_cases.json` و `contract_test_results.csv` و `contract_summary.csv` و `execution_meta.json`. تساوي النتائج المفردة 16 عقدًا × 4 تطبيقات = 64 تشغيلًا. وفي كل صف سجّل البرنامج النتيجة المتوقعة والملاحظة وحكم الاتفاق، لا «نجاحًا في منع حادثة».

التطبيق الاصطناعي	سياسة المدخلات (من 4)	ملف الخصائص (من 4)	النسخ (من 4)	إعادة المحاولة (من 4)	المجموع (من 16)
بلا حواجز	1	2	2	2	7
فحص المدخلات فقط	4	2	4	2	12
حدود الموارد فقط	1	4	2	4	11
الحواجز مجتمعة	4	4	4	4	16

نتيجة كل عقد اصطناعي في التطبيقات الأربعة

القيمة الثنائية تعني موافقة الناتج الفعلي لقاضي الاختبار المعلن

	بيانات السياسة				ملف الخصائص				ثقة النسخة				حد المحاولة				المجموع
	Q1	Q2	Q3	Q4	F1	F2	F3	F4	V1	V2	V3	V4	R1	R2	R3	R4	
بلا حواجز	1	0	0	0	1	1	0	0	1	1	0	0	1	1	0	0	7/16
حراسة المدخلات	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0	12/16
حراسة السعة	1	0	0	0	1	1	1	1	1	1	0	0	1	1	1	1	11/16
حواجز مجتمعة	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	16/16

1 = موافق 0 = غير موافق

16 حالة مصممة بدويًا × 4 تطبيقات اصطناعية = 64 تشغيلًا فعليًا. النتيجة اتفاق مع القاضي، وليست منعًا مقاسًا لأعطال إنتاجية.

مصفوفة نتائج 64 تشغيلًا على بيانات اصطناعية

شكل 1. المخرجات المحسوبة من `data/contract_summary.csv`. الدرجة مقدار مطابقة مخرجات مُصمَّمة يدويًا مع قاعدة حكم صُمِّمت معها؛ ليست قياسًا لفاعلية سوقية أو احتمالات حوادث حقيقية.

يتضح أن كل مسار تحقق جزئي نجح في عائلاتِ صُمِّم لمعالجتها وبقي ناقصًا في أخرى. لكن النتائج شبه محددة سلفًا؛ لا دلالة اختبار فروض إحصائي أو قيمة تعميم خارج هذا البرنامج. فائدتها تدريبية ومعمارية: توحيد لغة السيناريو، وفصل قاعدة الحكم عن آلية التنفيذ، والاحتفاظ بنتيجة كل عقد بصورة يمكن مقارنتها وإعادة توليدها. لا يمكن تفسير نجاح 16/16 على أنه نظام لا يتعطل.

6. الضبط والتحقق والحساسية: ما نُفذ وما لم يُنفذ

نُفذ: استعمال قواعد ثابتة بلا عشوائية، وتصدير جميع الحالات الأربع والستين، والتحقق برمجياً من عدد العقود ومن مجموع كل تطبيق، وتسجيل بصمة محتوى الحالات. ويمكن لمستخدم تشغيل الشيفرة في بيئة Python 3.10+ دون تبعيات خارجية، ومقارنة الملفات بعد إعادة التشغيل. هذه قابلية إعادة إنتاج حاسوبي (Computational Reproducibility) للمثال وليست إعادة لواقعة طبيعية.

لم يُنفذ: عكس العقود على أنظمة إنتاج، أو مراجعة استقلالية لصحة القاضي، أو قياس زمن تنفيذ فريق غير المؤلف، أو تحليل قبول عميل، أو اختبار اختراق، أو تقييم نموذج لغوي فعلي، أو إعادة تشغيل عدد كبير من بذور اصطناعية، أو اختبار تغيرات بيانات ذاتية التكيف. ليست هناك فواصل ثقة أو قيم احتمالية لهذا المثال، لأن المجموعة ليست عينة إحصائية. وجود 64 تشغيلاً لا يخلق مجتمعاً من 64 حادثة مستقلة.

حساسيات مهمة ينبغي إضافتها: ماذا إن تغيرت دلالة إزالة التكرار من مجموعة إلى قائمة ترتيب؟ ماذا إن كان تحديث قديم مسموحاً ضمن إجراء تراجع موثق؟ ماذا إن كانت إعادة المحاولة الأكثر من أربع مرات ضرورة تجارية في خدمة لا تؤذي مواردها؟ ماذا إن كانت الاستثناءات غير الممسوكة مرغوبة لتجنب سماح غير آمن؟ في هذه الحالات قد ينتقل جزء من التنفيذ الرابع من «صحيح» إلى «مخالف» بمجرد تغيير وظيفة النظام. وهذا بالضبط سبب الحاجة إلى معايرة دلالية مستقلة عند كل نقل.

7. الفجوة بين اختبار يعمل واختبار مفيد

لأن السيناريو يبدأ من خبر حقيقي، قد تُمنح النتيجة سلطة لا تستحقها. نصّ حادثة ليس مواصفة كاملة للبرامج التي تسببت فيها. قد يحذف المشغل معلومات لحماية نظامه، وقد تختلف وظيفة الحقول بين أدوات متشابهة ظاهرياً، وقد يستند التحقيق إلى أدلة جزئية. لهذا نميز في أي تجربة مستقبلية بين أربع نتائج محتملة: عطل أصلي جرى تمثيله بدقة، وعائلة آلية جرى تمثيلها تجريبياً، ومثال تعليمي يحاكي لفظ الآلية فحسب، وتجربة لا تستوفي شروط التطبيق ويجب عدم تشغيلها.

المثال الحالي يقع في الفئتين الثانية والثالثة بحسب العقد؛ لا يوجد أي تطابق مثبت مع المنظومات التاريخية. ويمكن أن تكون نتيجة PASS صحيحة بالنسبة لعقد غير صحيح. في الأعمال المعتمدة على الذكاء الاصطناعي تظهر مشكلة إضافية: قد يولد النموذج اللغوي تجربة سليمة النحو، ثم يملأ فراغات التقرير بعلاقة لم تُذكر فيه. قد يختلق النسبة أو اسم الحقل أو معنى التكرار. لذلك يُشترط فصل دقة الاستناد للمصدر (Evidence Grounding) عن صلاحية ترجمة العقد (Compilation Validity) وصحة دلالاته (Oracle Validity).

7.1 اختبارات جيدة يجب أن تفشل فيها الأداة نفسها

على المنتج أن يمتنع عن إصدار تجربة عندما يجهل الوظيفة. نقترح شواهد معايرة سالبة:

- نصّ حادثة يخفي تعريف الحقل اللازم، فيجب أن يطلب تحقيقًا إضافيًا.
 - تقرير فيه تفسير سببي يتراجع الناشر عنه، فيجب أن يسحب العقد أو يراجع نسخته.
 - خدمة تتعامل مع القوائم المرتبة، فيجب ألا تُزيل التكرار لمجرد نجاح حالة مشابهة في تطبيق آخر.
 - بيئة لا تتوافر فيها حدود إيقاف أو استرجاع مناسبة، فيجب ألا تصدر تجربة تنفيذية.
 - سجل مراقبة ناقص لا يثبت نتيجة القاضي، فيجب تصنيف الاختبار «غير حاسم» لا «ناجح».
- تقييم عدد الاختبارات المولدة فحسب قد يكافئ منصة تتحدث كثيرًا ولا تختبر شيئًا. المقياس الأقرب للمعنى هو عدد العقود المستقلة التي اجتازت فحص المصدر والدلالة والتنفيذ والنتيجة، ثم بقيت صالحة بعد تعديل نسخة العميل.

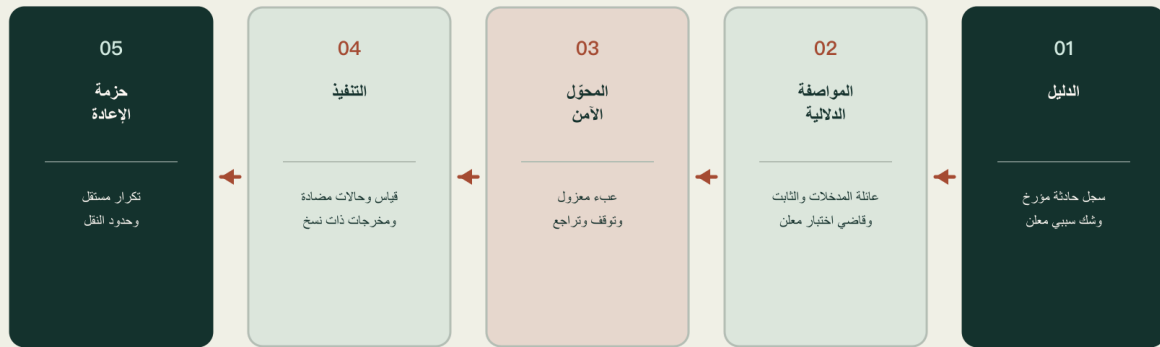
8. البنية الهندسية الأولية للمنتج

يمكن اختبار مفهوم مستقل في سبع مكونات، على أن تكون الحدود بينها واضحة:

- أ. مخزن تقارير بإصدارات (Versioned Incident Corpus).** يحتفظ بالرابط والتاريخ ونوع الدليل والنسخ والتصحيحات. لا يعيد توزيع محتوى محمي بلا إذن؛ يخزن بيانات منشأ ومقاطع مسموحة أو دلائل استرجاع.
- ب. مستخرج فرضيات (Hypothesis Extractor).** نموذج لغوي محلي أو متاح قانونيًا يرشّح الآليات ويفصل الملاحظات عن الاستنتاج، ويصرح بدرجة عدم الاكتمال. لا تُقبل تعليمات تظهر داخل تقرير مستورد على أنها أوامر للمنظومة.
- ج. محلّ صلاحية الدلالة (Semantic Applicability Checker).** يسأل عن وظائف الملف وحقوق التراجع والنسخ والحالة المشتركة، وهل يمكن تعريف قاعدة حكم مناسبة. قد يقترح أسئلة على مهندس العميل.
- د. مترجم تمثيل وسيط (Intermediate Representation Compiler).** يحول العقد المعتمد إلى مجموعات حالات وبرامج إدخال أعطال مقيدة. يفصل نوع الحالة عن أداة التنفيذ، ويحافظ على معلومات لا يستطيع تنفيذها في حقل unsupported بدل إسقاطها.
- هـ. مشغلات اختبار محدودة (Bounded Execution Adapters).** تعمل أولاً في بيئة محاكاة أو اختبار معزولة، بصلاحيات ضيقة وسقف للمدة والحمولة وآليات توقف وتراجع. لا تفترض توفر تلك الضمانات بصورة مطلقة لأن أدوات السوق نفسها تحتاج ضبطًا.
- و. قاضٍ وتحليل آثار (Oracle and Trace Analyzer).** يقارن المخرجات بخصائص معلنة، ويشير إلى أثر مفقود أو سلوك غير حاسم. يستفيد عند الملاءمة من «أوبن تليميتري» (S11 (OpenTelemetry مع عدم اختزال سلامة المعاملة في معدلات طلبات ناجحة.
- ز. حزمة شواهد وإعادة تشغيل (Evidence Replay Bundle).** توضح نسخة النظام والمدخلات والحكم والمحاذير والقيود. ويمكن لفريق مشتري مستقل إعادة التجربة على مختبره أو رفضها إذا اختلفت شروطه.

السلسلة المقترحة من الحادثة إلى الاختبار

خمس ترجمات معرفية؛ كل مرحلة تحتاج دليلاً وتحققاً خاصاً بالتطبيق



بنية مفاهيمية غير منفذة كمنصة. يجب حفظ عدم اليقين السببي، واختبار الدلالة، وتقييد التشغيل. ثم يجب تمكين مشترٍ مستقل من إعادة النتيجة أو رفض النقل.

مسار تحويل سجل الحادثة إلى منتج اختبار قابل للمراجعة

شكل 2. سلسلة إنتاج مفاهيمية. المراحل ليست نظاماً سعودياً قائماً أو معياراً منشوراً، ويجب التحقق من كل واجهة وتقييمها قبل استخدامها على بيئات حقيقية.

المنافسة المحتملة ليست إنشاء كل هذه الأجزاء من الصفر. يمكن بناء محوّلات فوق واجهات قائمة، بما فيها أدوات تجارب الأعطال ومشاريع الرصد المفتوحة؛ القيمة المميزة التي تحتاج إثباتاً هي انخفاض تكلفة نقل المعنى ووضوح دليل الاختبار لدى عملاء مختلفين. ولأن وظيفة المنصات لا تنحصر في السحابة، قد تُحتاج وحدات دون اتصال بشبكة العملاء، أو مكتبات توليد مدخلات لبرامج حاوياتهم، لكن هذا قرار هندسي حسب حالة الشراء لا شعار تصميم.

9. اقتصاد المنتج: من التجربة الأولى إلى العميل الثالث

نعرّف ثلاثة نماذج عمل منافسة على نحو وصفي:

1. استشارة اختبار مخصصة (Bespoke Reliability Consulting). يشتري العميل وقت خبراء يعاينون نظامه ويصممون تجربة خاصة. قد يكون الإيراد مرتفعاً لكل عقد، لكن يتكرر معظم الجهد، وتعتمد الهوامش على ساعات المهندس.
2. حزمة سيناريوهات مرخصة (Licensed Testpack). يشتري العميل مكتبة يعيد تشغيلها وصيانتها مع تعديل محدود؛ يحتاج البائع إلى أدوات تكيف ومعايير نسخ وصلاحية أو يحدث تضخم طلبات الدعم.
3. خدمة تشغيل وشواهد متكررة (Subscription Execution Evidence). يدفع العميل لربط تجارب بنسخ متغيرة باستمرار؛ توفر الاشتراكات إيراداً متكرراً إذا حققت قيمة تتجدد، لكنها تخلق مسؤوليات تشغيل وتخزين وأمن وخصوصية.

لتجنب إعلان سوق غير مثبتة، نقترح قياسات لوحدة المنتج لكل عميل أجنبي تجريبي:

الرمز	المتغير	قياسه المقترح
H_i	ساعات التهيئة الأولى	سجلات عمل محددة تربط العمل بالعميل والنسخة
H_r	ساعات تحديث المنتج عند نسخة جديدة	جهد ترقية لا يتضمن أعمالاً تجارية لا علاقة لها بالمنتج
H_c	ساعات دعم خاص بالعميل	وقت تدخل لا يمكن تعميمه على مشتريين آخرين
P	سعر مدفوع	إيراد تعاقدى متحقق، لا قائمة أسعار أو تفاهم مبدئي
V	منفعة موثقة	حالات إضافية ذات صلة لم تكشفها مرجعية العميل، مع تكاليف تنفيذها
R	التجديد المدفوع	تمديد فعلي بعد مراجعة نتائج النسخة الأولى

يمكن تعريف نسبة جهد قابل لإعادة الاستخدام (Reusable Engineering Share) بشكل مبدئي كالتالي:

$$\mathcal{S} = 1 - \frac{H_c}{H_i + H_r + H_c}$$

إذا كانت الساعات جميعها غير سالبة والمقام موجباً، تقع النسبة بين صفر وواحد. لكنها ليست هامش الربح (Gross Margin)، ولا تقيس وحدها جودة المنتج. فارتفاعها قد يدل على منصة جيدة، وقد يدل على عدم تقديم دعم احتاجه العملاء فعلاً. وتحتاج المقارنة إلى تعريف متسق للفترة والنظام والمشتري. لا أرقام مقيسة لهذه النسبة في الدراسة الحالية.

كما يمكن للمشتري تكوين حدٍّ أعلى افتراضي للقيمة المتوقعة عبر احتمال عطل وخسارته وتغير احتمال تجنبه بفضل الاختبار، ثم خصم تكلفة التجربة والسعر والآثار الجانبية. لكن تحديد الأثر السببي غير مباشر: فقد تتبع نجاح الاختبار إعادة تصميم شاملة، وقد تنخفض الأعطال بسبب تحسين التدريب أو انحسار الطلب أو تغير البنية؛ وقد لا يقع العطل أصلاً خلال عمر العقد. لهذا لا نستنتج عائد استثمار من قصص العملاء أو من عدد العقود المنفذة.

10. كيف نفرق المنتج القابل للبيع من الخدمة المتخفية في هيئة منصة؟

التمييز العملي يحتاج لوحة قياس منشورة داخل المشروع، لا ألقاب تسويقية. نتوقع من منتج برمجي قابل للتوسع أن يقل متوسط جهد التخصيص عند الانتقال إلى عميل جديد يعمل على بنية مشابهة، وأن يحتفظ بنتيجة مفيدة بعد تحديثات منتج العميل، وأن يستطيع مهندس من جهة مشتري التشغيل دون حضور مطوري المنتج كل مرة. أما الخدمة المتخصصة فقد تنجح اقتصادياً مع بقاء هذه الخصائص ضعيفة، وهذا مشروع آخر له حساب مختلف.

نقترح تجربة سوق تبدأ بقطاعات مدنية يمكن الحصول فيها على إذن واضح للاختبار: برمجيات أعمال متعددة المستأجرين، ومنصات معالجة طلبات غير حساسة في بيئات ما قبل الإنتاج، ومكونات عامة لإدارة التكوين. وعند إثبات النجاح يمكن بحث فئات أكثر حساسية مع الضوابط اللازمة. لا نرى أن نشر تفاصيل محاكاة منشآت حيوية أو ربط المنتج ببيانات تشغيل خاصة شرط للانتقال إلى السوق.

يشكل معيار دخول أولي صارم، على سبيل بروتوكول مستقبلي مقترح وليس نتيجة، ثلاثة عملاء دوليين مستقلين على منصتين مختلفتين، ينجح واحد منهم على الأقل في إعادة تشغيل حزمة اختبار دون تدخل الفريق البائع، مع تسجيل تكلفة التهيئة والملاحظات والحالات التي فشلت في النقل. هذه العتبة استدلائية للتحفيز على اختبار قابل للدحض، وليست قاعدة علمية منشورة لتصنيف الشركات.

11. تصميم الدراسة التي قد تدحض الفكرة

11.1 العينة والمقارنات

يُستخرج سجل أولي من حوادث منشورة أصلية، مع تصنيف مسبق للمصادر حسب استقلاليته وتفصيلها. يجب فصل مؤسسات التدريب عن مؤسسات الاختبار، وألا تدخل في مجموعة التدريب حادثة من نفس عائلة الحالة المحجوزة إذا كنا نريد قياس القدرة على التعميم. ينبغي كذلك فصل الحوادث بحسب تاريخ نشرها، لا السماح لنموذج لغوي مدرب على سجل أحدث باستخدام معرفة المستقبل في اختبار يزعم أنها لم تكن متاحة وقت إصدار عقده.

المقارنة الأولى مع مهندس يكتب اختبارًا يدويًا باستخدام أدواته الحالية، والثانية مع مكتبة سيناريوهات جاهزة من مقدّم الخدمة السحابية، والثالثة مع توليد قالب تجريبي بلغة طبيعية دون فحص دلالي إضافي، كما يظهر في بعض ممارسات السوق S06S07. يجب أن يتساوى الوقت المتاح للفرق قدر الإمكان وأن تسجل مهارة المهندسين، وإلا تحول تفوق فريق إلى تفوق أداة.

11.2 المخرجات ومعايير القبول

- **صحة منشأ الآلية (Provenance Accuracy):** هل تُسند الآلية إلى موضع محدد في تقرير أصلي ولم تُضف تفاصيل غير موجودة؟
- **اتفاق القضاة (Oracle Agreement):** هل يتفق خبراء مستقلون، بعد اطلاعهم على وظيفة النظام، مع النتيجة المتوقعة؟ تُسجل حالات الاختلاف لا تُحذف.
- **صلاحية التطبيق (Applicability Precision):** ما نسبة العقود التي حُكم بصلاحيته والتي يتبين عند التحقق أنها صالحة حقًا؟
- **مردود النقل (Transfer Yield):** نسبة العقود التي بقيت صحيحة الدلالة بعد نقلها إلى بيئة جديدة ضمن العينة المحددة، مع الإفصاح عن المقام والاستثناءات.

- **المنفعة الإضافية (Incremental Fault Discovery):** عدد الاختلالات الصحيحة الإضافية التي كشفها المنتج مقارنة بالأداة الأصلية، بعد استبعاد التكرار والاكتشافات الزائفة.
 - **تكلفة التوسع (Marginal Integration Cost):** ساعات التطوير والتوثيق والتشغيل لدى فريق البائع والعميل لكل بيئة ونسخة جديدة.
 - **الطلب الخارجي المتكرر (Foreign Paid Retention):** عدد المشتريين الأجانب الذين جددوا بعد دورة استخدام فعلية، والقيمة المدفوعة الصافية، وليس عدد العروض أو مذكرات التفاهم.
- عتبة التأكيد:** إذا لم يُحسن المنتج اكتشاف اختلالات وظيفية في اختبار أعمى مقارنة بأدوات قائمة ضمن الوقت والتكلفة نفسه، أو احتاج في كل مشروع جهد استشارة لا يتراجع، أو امتنع المشترون الأجانب عن التجديد، تضعف فرضية الصناعة البرمجية التصديرية المستقلة. قد تظل هناك قيمة تعليمية أو استشارية رغم ذلك.

11.3 الاستدلال السببي

حتى لو انخفضت حوادث عميل بعد اقتناء الحزمة، لا يجوز نسب الانخفاض إليها من دون مجموعة مقارنة أو تصميم تحليلي يُعالج تغير الإصدارات وشدة الاستخدام وأعمال الإصلاح الأخرى. المخرج الذي يملك أقوى اتصال مباشر بالمنتج في البداية هو كشف اختلال جديد وقابل لإعادة التشغيل، أو تقليل زمن الوصول إلى اختبار موثوق؛ أما منع حادثة في المستقبل فيحتاج رصدًا أطول وتفسيرًا مضادًا للواقع. وللأعطال الشديدة ندرة تجعل إثبات فائدة انخفاضها إحصائيًا صعبًا، حتى لو كانت قيمتها الاقتصادية كبيرة.

12. الصلاحيات والأمان بوصفهما خصائص تنفيذ، لا شعارين إداريين

التجربة التي تكشف فشلًا لا تستحق النجاح العلمي إذا سببت تعطلًا حقيقيًا خارج النطاق المصرح به. ينبغي أن تبدأ كل الحزم في بيئة معزولة أو ما قبل الإنتاج، وتطلب موافقة نطاق واضحة إذا استعملت حقنًا في بيئة حية، وتضع حدودًا للمدة والحمل والموارد وقواعد توقف ترتبط بمؤشر صالح. توثق «أمازون» (AWS) شروط توقف للتجارب مبنية على إنذارات مراقبة محددة S08، وهو مثال قائم على أن السلامة عنصر من تصميم التجربة وليس عبارة في نهايتها.

لا ينبغي حفظ كلمات مرور أو سجلات اتصال أو خرائط تبعيات دقيقة لمرافق حساسة في الحزم القابلة للبيع أو النشر. ويجب فصل أسماء الحوادث العامة من بيانات تشغيل العميل ومراعاة حقوق المحتوى والتراخيص عند استخراج معلومات من تقارير منشورة. إن بناء تصنيف عالمي للأخطاء لا يحتاج معرفة القاعدة السرية التي استعملتها جهة ما أو نشر حسابها الداخلي. واستخدام الذكاء الاصطناعي ينبغي أن يلتزم بالعزل بين مادة المصدر غير الموثوقة والتعليمات التشغيلية، وألا يتحول استخراج نص إلى سلطة تنفيذ.

13. ماذا تضيف هذه الدراسة إلى الأسئلة السعودية حتى 2035؟

تضيف الدراسة تمييزًا بين أربع درجات من القيمة:

1. **الخبرة المحلية (Local Expertise):** تستطيع فرق سعودية إصلاح مشكلات حقيقية؛ وهذا مهم لكنه لا يكفي للتصدير.
 2. **المعرفة القابلة للتدوين (Codified Knowledge):** تُترجم الخبرة إلى مواصفات ودلائل قابلة للفحص؛ وهذا لا يثبت قابلية النقل.
 3. **المنتج القابل للنقل (Transferable Product):** يحقق الاختبار معنىً متقاربًا لدى عميل جديد بتكلفة محدودة؛ وهذا لا يثبت استدامة الطلب.
 4. **الصناعة التصديرية المتكررة (Repeatable Export Industry):** يدفع مشتررون مستقلون ويجددون ويستطيع المنتج تمويل صيافته وتحسينه؛ وهذه نتيجة اقتصادية تحتاج دليلًا مختلفًا عن البرهان الهندسي.
- هذه درجات وصفية مقترحة للتفكير وليست مؤشرًا رسميًا. ومن الخطأ أن يُستدل بالمستوى الأول مباشرة على الرابع، أو أن تُستعمل مشاركة 28 شركة في معرض دولي بدل معرفة حجم الإيراد التقني المستدام S04.
- الاستثمار البحثي السعودي الأكثر جدوى ضمن هذه الفرضية قد يكون مختبرًا محدودًا لا شركة خدمات شاملة: مجموعة حوادث موثقة متنوعة، وبروتوكول لترميز شروطها، ومُحوّلات مفتوحة لتجارب معزولة، ومسابقات تعمية بين حلول يكتبها البشر وحلول يقترحها نموذج لغوي ثم يراجعها مهندسون، ومعايير نشر تمنع ادعاءات قدرة لا تدعمها النتائج. إذا نجح ذلك تحت قياس مستقل، يمكن التفكير في منتجات متخصصة وعملاء أجنب. وإذا فشل، ينبغي نشر الفشل؛ لأنه هو نفسه مادة البحث التي تريد الصناعة أن تتعلم منها.

14. حدود الدراسة ونهاية تكليفها

النتيجة المنجزة الوحيدة من ناحية الحساب هي مطابقة 64 مخرجًا لعقود اصطناعية محددة مسبقًا. والمراجعة الخارجية محددة بمصادر أصلية منتقاة؛ ليست مراجعة منهجية شاملة (Systematic Review)، ولم تُجر عملية فحص منشورات مفهومة بقواعد استبعاد مسبقًا. ولم نحسب حجم سوق أو سعر عقد أو احتمال حادثة أو عائد استثمار حقيقي. ولم نتحقق من أن النموذج المفاهيمي يمتلك جودة علمية مطلقة، ولا أنه سينتج براءة اختراع، ولا أنه مُستقل عن الأدوات الموجودة.

الاستنتاج المقبول أضيق ولكنه جاد: من الممكن فنيًا وصف عائلات محدودة من إخفاقات إدخال البيانات والتحديث والتحميل في عقود قابلة للتشغيل وإخراج شواهد قابلة للفحص، كما يبرهن مثالنا الصغير. وتوجد في السوق أدوات مستقرة تمكّن صوّراً قريبة من هندسة الأعطال وتوليدها باللغة الطبيعية؛ لذا ينبغي اختبار القيمة المضافة المفترضة في صحة النقل الدلالي وانخفاض تكلفة التخصيص وتكرار الدفع الدولي. من دون تلك الاختبارات تبقى الصناعة المقترحة فرصة للبحث، لا خبرًا عن اقتصاد سعودي جديد.

15. التوافر والتكرار

البيانات المعروضة في ملفات data / اصطناعية بالكامل ومنشأة لهذا العمل، وهي مرخصة هنا بوصفها أمثلة تعليمية داخل الحزمة لا بيانات لجهة ثالثة. تُحفظ الشيفرة في analysis/reproduce.py بلا تبعيات غير مكتبة «بايثون» القياسية (Python Standard Library). شغلها من المجلد الرئيسي للبحث لتحصل على الملفات نفسها مع بصمة حالات الاختبار. تُقدّم الرسوم المستخرجة بصيغة متجهة (SVG) وصورة (PNG)، وجدول بيانات (Excel) للمراجعة البصرية والصيغ الحسابية المعلنة. لا تحتاج إعادة التشغيل إلى خدمات سحابية أو تسجيل دخول أو بيانات مخفية.

لا تتضمن الحزمة بيانات تشغيل سعودية أصلية، أو معلومات أسرار شركات، أو شيفرات هجوم على مرافق، أو تقارير نزاع لم يمكن التحقق منها. والاقتباسات من التقارير الأصلية في هذه الورقة بالمعنى، لا بنصوص مطولة أو بيانات تنفيذ حساسة.

المراجع

1. Cloudflare, Matthew Prince, *Cloudflare outage on November 18, 2025*, original incident postmortem, 18 November 2025. <https://blog.cloudflare.com/18-november-2025-outage>
2. Google Cloud, *Incident details: June 12, 2025 service disruption*, incident status and root-cause note, 2025. <https://status.cloud.google.com/incidents/ow5i3PPK96RduMcb1SsW>
3. Federal Communications Commission (FCC), *February 22, 2024 AT&T Mobility Network Outage*, Public Safety and Homeland Security Bureau Report, 22 July 2024. <https://docs.fcc.gov/public/attachments/DOC-404150A1.pdf>
4. Saudi Press Agency (SPA), *'Made in Saudi' Pavilion at GITEX AFRICA Showcases 28 Leading Tech Companies*, 14 April 2025. <https://www.spa.gov.sa/en/N2297562>
5. Saudi Press Agency (SPA), *Elm Signs MoU with Huawei to Advance Cooperation in AI and Digital Infrastructure*, 9 July 2026. <https://www.spa.gov.sa/en/N2630981>
6. Amazon Web Services (AWS), *Bootstrap your chaos engineering journey with AWS Fault Injection Service Scenarios Library*, 1 October 2024. <https://aws.amazon.com/blogs/mt/bootstrap-your-chaos-engineering-journey-with-aws-fault-injection-service-scenarios-library>
7. Amazon Web Services (AWS), *Chaos engineering made clear: Generate AWS FIS experiments using natural language through Amazon Bedrock*, 13 May 2025.

<https://aws.amazon.com/blogs/publicsector/chaos-engineering-made-clear-generate-aws-fis-experiments-using-natural-language-through-amazon-bedrock>

Amazon Web Services (AWS), *Stop conditions for AWS FIS*, online technical documentation. <https://docs.aws.amazon.com/fis/latest/userguide/stop-conditions.html>

Jepsen (Kyle Kingsbury / Aphyr), *Analyses and Research Ethics*, technical analyses and methods. <https://jepsen.io/analyses> and <https://jepsen.io/analyses/ethics>

Principles of Chaos Engineering, *Principles of Chaos Engineering*, foundational /practitioner document. <https://principlesofchaos.org>

Cloud Native Computing Foundation (CNCF), *OpenTelemetry has graduated... Now what?*, 24 July 2026. <https://www.cncf.io/blog/2026/07/24/opentelemetry-has-graduated-now-what>

Amazon Web Services (AWS), *REL12-BP04 Test resiliency using chaos engineering*, Well-Architected Reliability Pillar. https://docs.aws.amazon.com/wellarchitected/latest/reliability-pillar/rel_testing_resiliency_failure_injection_resiliency.html

Author-supplied materials, *Executable Incident Contract Demonstration*, version 1.0. Python code and CSV/JSON outputs are included in this package; explicitly .synthetic, not independently peer reviewed