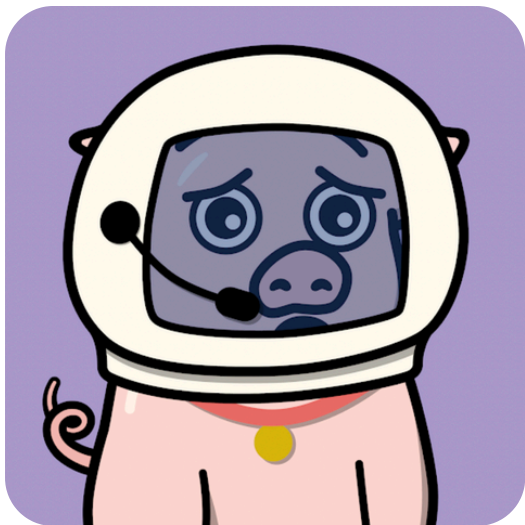


Gemma Meetup 2026 [Road to DevFest Tokyo #1] 2026/10/01

PCで動く Gemma 4による 自然言語テスト自動化

中山 慶 株式会社CureApp GitHub: kexi · X: @keisan

自己紹介



中山 慶

Kei Nakayama (kexi)

株式会社CureApp

技術部

GitHub: [kexi](#) X: [@keisan](#)

作ったもの：gemma-e2e-android

自然言語のゴールから、PC で動く Gemma 4 が画面の UI ツリーを読んで操作・判定する E2E テストツール。

gemma-e2e

SCENARIOS

a11y

auth

lab

regression

shop

smoke

web

Accessibility lab, ...

1 case ·

http://localhost:5174/?lab=a11y

a11y

lab

Pages through the lab screens with all personas

default model

max 12

Accessibility lab, ...

1 case ·

dev.xexi.gemmae2e.example/.MainActivity

a11y

lab

Pages through the lab screens with all personas

← DEVICE

Coffee shop (Android)

passed

shop · 2 cases · started 9/28/2026, 2:13:06 PM

✓

 all 2 cases passed

passed

 Adds a bean to the cart and sees the total

gemma-4-26b-a4b-it-qat-mlx

 6 steps

passed

 Places an order using the confirmation code

gemma-4-26b-a4b-it-qat-mlx

 12 steps

Check that a customer can complete an order. Log in with the email "demo@example.com" and the password "demo1234", add any one coffee bean to the cart, and go through checkout. The checkout screen shows a confirmation code; the screen after it asks you to type that code back in. Enter it, place the order, and confirm that an order number is shown.

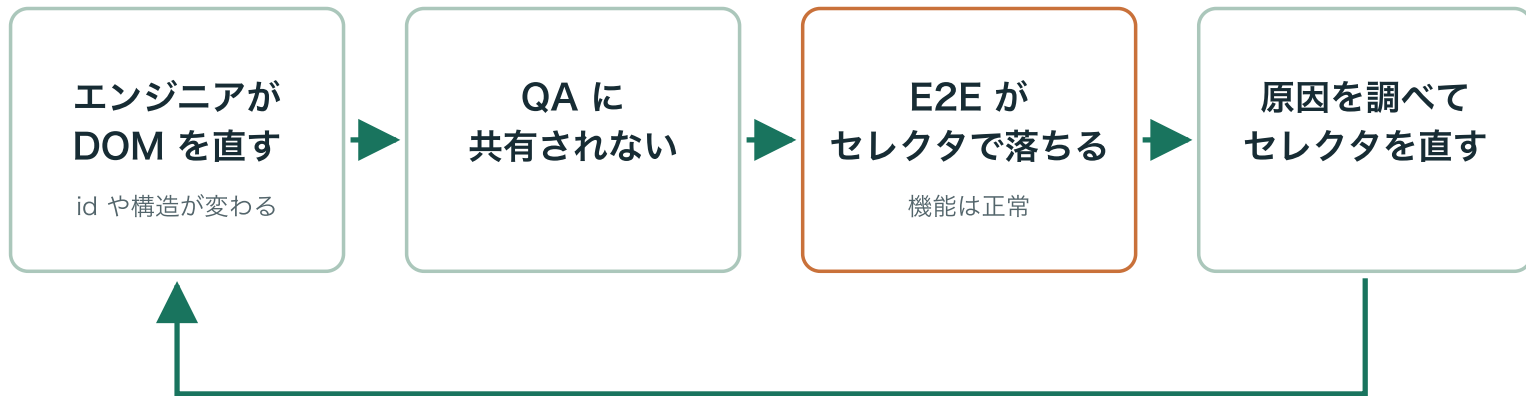
✓

 The order was placed successfully and the order number 'KCS-1001' is displayed on the screen.

操作の判断は Computer use ではない。UI ツリーのテキストを読んで選ぶ。
アプリのソースコードは見ない。手がかりは実行中の画面だけ。

作ったきっかけ

① セレクタの修正に疲れた



UI を変えるたびに繰り返す

作ったきっかけ

② ペルソナごとのテストは、コストが高くてできていなかった

ペルソナ

×

画面

×

リリースのたび

色の見分けにくさや老眼など、見え方の違う人ごとに画面を確かめたい。でも人手では回せない。

テストに書くのは、達成したいこと

ログインして、コーヒー豆を 1 つ注文する。

途中で表示される確認コードを次の画面に入力し、

注文番号が表示されることを確認する。

この文章をそのまま E2E テストとして流す。セレクトも手順も書かない。

画面を読み、1 手ずつ操作を繰り返す



Android は ADB、Web は CDP で動かす

Android : ADB

Android Debug Bridge

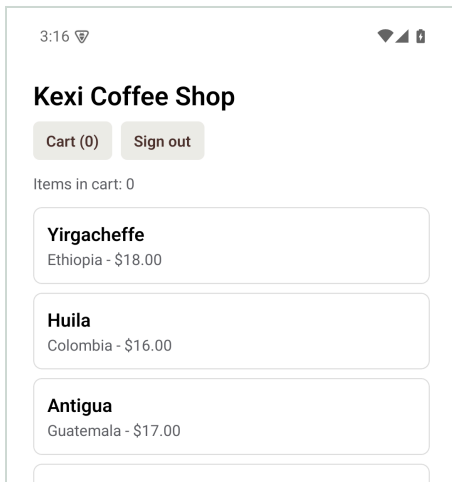
PC から端末につなぎ、画面の取得やタップ・入力のコマンドを送る公式ツール。

Web : CDP

Chrome DevTools Protocol

外部のプログラムから Chrome を操作し、DOM を読むためのプロトコル。

画面の UI ツリーを、XML で取り出す

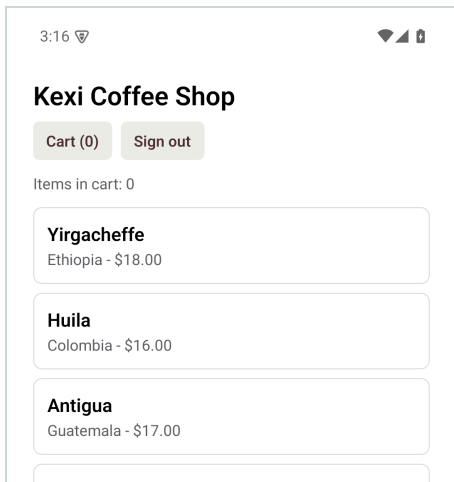


Android 標準の UI ツリー出力（uiautomator dump）の XML

```
<node index="3" text="" resource-id="signOutButton"
  class="android.widget.Button"
  content-desc="Sign out button"
  clickable="true" enabled="true" focused="false"
  bounds="[274,274][475,367]" ...>
  <node text="Sign out"
    class="android.widget.TextView" ... />
</node>
```

Gemma が読む、画面のテキスト

XML のままだと精度が落ちたので、ノイズを減らしてテキストに変換し、Gemma へ渡す。



変換したテキスト → Gemma へ

```
FrameLayout
```

```
TextView text="Kexi Coffee Shop" id=screenTitle
```

```
[0] Button desc="Cart button" id=cartButton
```

```
TextView text="Cart (0)"
```

```
[1] Button desc="Sign out button" id=signOutButton
```

```
TextView text="Sign out"
```

```
TextView text="Items in cart: 0" id=cartCount
```

```
ViewGroup
```

```
[2] Button desc="Yirgacheffe row" id=beanRow-yirgacheffe
```

```
...
```

Gemma は毎回まっさらなので、経緯を渡す

呼び出しのたびに、前のことは覚えていない。だから毎ステップ、次の 4 つを渡し直す。

ゴール

シナリオの文章

メモ

画面で見た値を Gemma が
書き残したもの

これまでの操作

直近 30 手

今の画面

取り直した UI ツリー

7 手目

画面：確認コード 4821 が表示されている

Gemma：4821 をメモに残す



10 手目

画面：確認コードの入力欄

(4821 はもう表示されていない)

Gemma：メモにある 4821 を入力

ペルソナごとに、画面の見え方もレビューする

画面 → Gemma（画像） × ペルソナ → 場所・理由・提案

ペルソナ	仕込んだ問題	Gemma の指摘（抜粋）
赤・緑の見分けにくさ	在庫を赤・緑の点だけで示す	✅ 「在庫状況が色（緑と赤）のみで示されており」
青・黄の見分けにくさ	ホット/アイスを黄・青の四角だけで示す	✅ 「区別が、色（黄色と青色の四角形）のみで…」
老眼	送料の注釈を小さい文字で書く	✅ 「文字が小さく、…読みづらい」
弱視	送料の注釈を薄い灰色で書く	✅ 「文字の色が薄いグレー…コントラストが低く」
弱視	小さなアイコンを密集させる	✅ 「アイコンが密集して…視覚的な混雑」

Gemma 4 26B-A4B QAT（MLX 4bit、thinking mode）、Web の検証用 6 画面、1 ケース 1 ペルソナ、各 1 回。

指摘の例：青・黄の見分けにくさ

Brewing guide

 Best hot

 Best iced



 Yirgacheffe



 Huila



 Antigua



 Sidamo



 Geisha



 Toraja

Next

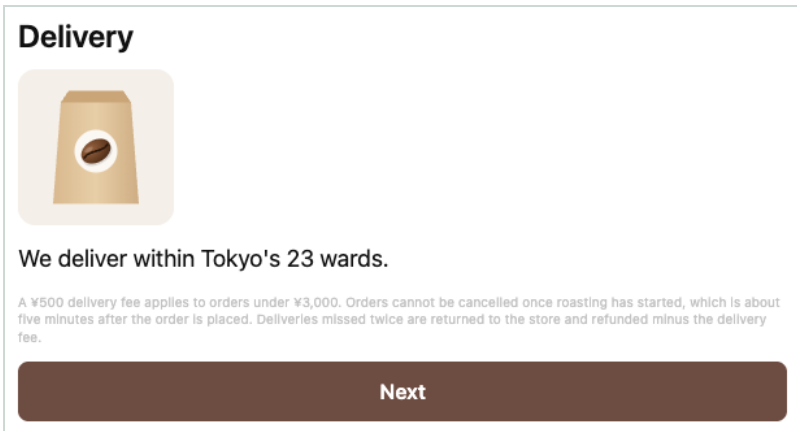
場所 「Best hot」と「Best iced」の凡例および各リスト項目

色だけで区別

理由 「Best hot」と「Best iced」の区別が、色（黄色と青色の四角形）のみで行われています。青と黄色の識別が困難なユーザーにとって、情報の区別ができません。

提案 色だけでなく、テキストラベル（例：「Hot」や「Iced」）を追加するか、異なる形状のアイコン（例：熱いものは太陽、冷たいものは雪の結晶）を使用してください。

指摘の例：老眼



場所 画面中央の注釈テキスト（A ¥500 delivery fee...）

文字の大きさ

理由 文字が小さく、老眼などの視覚特性を持つユーザーにとって読みづらい可能性があります。

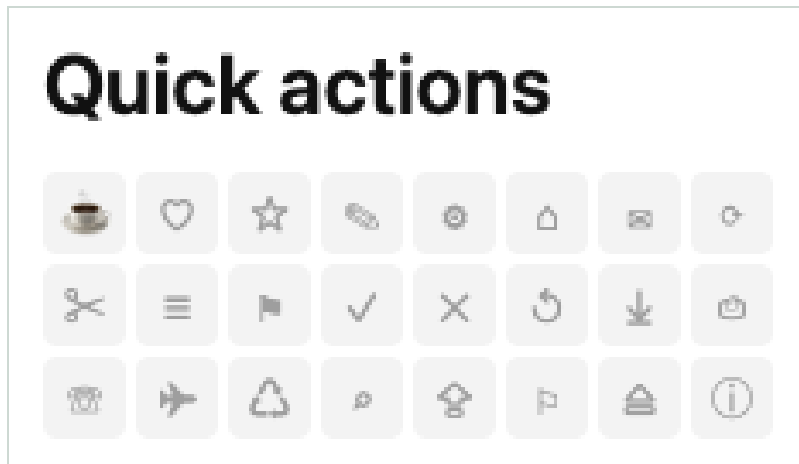
提案 文字サイズを大きくして、可読性を向上させてください。

コントラスト

理由 背景の白に対して文字の色が薄いグレーであり、コントラストが低いため、読み取りにくい可能性があります。

提案 文字の色をより濃い色に変更し、背景とのコントラストを高めてください。

指摘の例：弱視



アイコン部分を拡大

場所 「Quick actions」の下にあるアイコン一覧

小ささ

理由 アイコンが非常に小さく、線の太さも細かいため、詳細を識別しにくい。

提案 アイコンのサイズを大きくし、線の太さを太くして視認性を高める。

密集

理由 アイコンが密集して配置されており、視覚的な混雑が生じている。

提案 アイコン同士の間隔を広げ、個々の要素を明確に区別できるようにする。

ペルソナを広げる案

ペルソナは名前と説明文だけで定義する。見え方以外の困りごととも書ける。

漢字が読めない外国人

日本語の会話はできるが、漢字はほとんど読めない。ひらがな・カタカナ・英語は読める。

指摘の例

「注文を確定」が漢字だけで、押してよいか判断できない

スマホに不慣れな高齢者

スマホを使い始めたばかり。アイコンだけのボタンの意味を知らない。

指摘の例

カートのボタンがアイコンだけで、何のボタンかわからない

片手で操作する人

電車で立ったまま、親指だけで操作する。

指摘の例

「注文する」が画面の上端にあり、指が届きにくい

ローカル Gemma を使う理由

外部に流出しない

画面の内容・テスト用の認証情報・
スクリーンショットを、
外部の API に送らない

結果が安定する

重みが手元にあるので、
提供側の都合でモデルが
変わったり劣化したりしない

クラウドのモデルは、同じモデル・同じバージョンでも日によって調子が変わる。

Gemma 4 を使ってみて

十分に賢い。

- UI ツリーを解釈して、自然言語のゴールどおりに操作できた
- 画面の画像から、ペルソナごとの見え方の問題を指摘できた

Gemma に UI 操作を任せる。

テストの意図と根拠は、人が持つ。

ゴールから次の操作を選び、画面を読み直す。
UI 変更への追従性と、合否判定の信頼性を検証していく。

ソースコード

github.com/kexi/gemma-e2e-android