

ユーザビリティ指標 + フレームワーク

ゴール

プラットフォームのユーザビリティ指標を確立し、それによってさまざまな機能のユーザー体験を測定できるようにする。指標は、現在進行中の製品開発チームをサポートし、開発中の新機能がユーザーニーズを中心に据えていることを確認し、ユーザーコミュニティの成長を可能にする。

ユーザビリティ指標とフレームワークについて

私たちはここで挙げたユーザビリティ指標を選択しました。このうち、4つ（1～4）は、測定する高レベル指標です。

- プラットフォームの全体的な使いやすさ
- ユーザーによる宣伝の可能性
- ユーザーの満足度
- ユーザーの「安心感」と「プライバシー」

これらの4つの指標は、アンケートを通じてユーザーの回答から測定されます。

4つの指標（5～8）は、4つのパラメータにおけるセキュリティ機能を具体的に測定します。

- ユーザーの労力
- 成功
- 効率
- エラー率

機能を使用しているとき

これら4つの指標はユーザーインタビューを通じて測定されます。これらの指標はユーザビリティの定量的な見解を提供しますが、ユーザビリティテストは定性分析のために各機能ごとに実施され、ユーザビリティにおける問題点やギャップに基づく推奨事項や提案を提供します。このプロセスに関する詳細は、こちらをご覧ください。 [here](#)

指標の使い方

指標の使用方法 スコアを分析し、ユーザーに掘り下げた質問をして、スコアを改善する方法を見つけ出します。

この指標は拡張性があり、このプロジェクトの範囲を超える新機能がリリースされた場合でも、同様の調査+ユーザビリティ・テストの方法論を適用することができます。

ユーザビリティ指標

凡例:

👉 概要

📋 入力収集方法

🌐 ベンチマークスコア

🕒 頻度/測定タイミング

📐 測定方法

1. システムユーザビリティスコア (SUS)

👉 システムの全体的なユーザビリティのレベルを大まかに示す。このスコアは、何が機能し何が機能しないかを具体的に特定することなく、ユーザビリティの全体像を提供する。

📋 Survey - SUSスコアは、回答者が5段階の「強く同意する」から「強く同意しない」の尺度で評価する10の標準的なステートメントを提供するユーザー調査を通じて収集されます。

🌐 良いスコアは68以上とされる

68をベンチマークとして選択したことは、ある意味で恣意的なものであるが、調査と経験に基づいて、妥当な基準点であることが分かっている。SUSを開発したジョン・ブルックのオリジナルの研究では、68前後のスコアは「合格点」のユーザビリティであるとみなすことができると提案している。つまり、68以上のSUSスコアは一般的に、そのシステムがユーザビリティに優れていると認識されていることを示し、68以下のスコアは潜在的なユーザビリティの問題を示唆している。

🕒 四半期

📐 Calculator. SUSスコアは、まず各項目からのスコア寄与を合計することで算出されます。各項目のスコア寄与は0から4の範囲です。

1、3、5、7、9の各項目については、スコアへの寄与度はスケール位置から1を引いた値となります。2、4、6、8、10の各項目については、スコアへの寄与度はスケール位置から5を引いた値となります。各項目のスコアの合計に2.5を掛けると、SUSの総合値が得られます。SUSスコアは0から100までの範囲となります。

2. ネットプロモータースコア正味推奨者比率 (NPS)

👉 顧客満足度とロイヤリティを示す指標として、その製品/プラットフォームを他人に推薦する可能性を尋ねる。

[Survey](#) - 10段階評価で「まったくそう思わない」から「非常にそう思う」までで収集-100から+100の範囲で利用可能な場合、0以上のスコアは「良い」とみなされます。計算においては、推奨者と非推奨者のどちらに転ぶかという点が重要であるため、消極的な投票は無視されます。推奨者はスコアをプラスに、非推奨者はスコアをマイナスにします。非推奨者のユーザーの方が多い可能性があるため、スコアがマイナスになることもあります。そのため、ニュートラルな0より上であれば、すべて「良い」とみなされます。

🕒 四半期

👉 [Calculator](#). NPSは、NPSの質問に6以下で回答した顧客の割合（「批判者」と呼ばれる）から、9または10で回答した顧客の割合（「推奨者」と呼ばれる）を差し引いて算出されます。

3. ユーザー満足度 (USAS)

👉 ユーザーの満足度を示す

[Survey](#) - 1～5のスケールで収集：「製品/プラットフォームにどの程度満足しているか」 🌟 良いスコアは通常、75%から85%の間になります。合計スコアが40の場合、30～35の範囲のスコアは75%～87.5%に相当し、良いスコアと言えます。それ以上のスコアは素晴らしいスコアです。

🕒 四半期

👉 [Calculator](#). ユーザーには、体験のしやすさを1から5までの数値、または怒りから喜びまでの顔文字で評価してもらいます。収集した回答を平均化することで、特定のプロセスがユーザーにどの程度の負担を強いるかを把握することができます。

4. ユーザーセキュリティスコア (USS)

👉 ユーザーがどの程度安全で安心だと感じているかの表示

[Survey](#) - [TUSSスコア](#)は、回答者が5段階の「強く同意する」から「強く同意しない」の尺度で評価する8つのステートメントを提供するユーザー調査を通じて収集されます。 🌟 高いほど良い

🕒 四半期

👉 [Calculator](#). USSは、ユーザー回答の加重平均です。すなわち、ユーザーの努力評価の合計を回答総数で割った値です。

5. ユーザ労力スコア (UES)

👉 ユーザーが機能を利用するために費やさなければならなかった努力の量の指標

📋 Survey (pop-up, in-app) - - 例: 「[特定の機能]の操作は簡単でしたか？」

🎯 1-5または1-7のスケールで、高いほど良い (質問がタスクの実行のしやすさに関する場合)

🕒 機能リリース時

🔢 Calculator. UESは、やりとりが簡単だったと回答した顧客の総数を回答の総数で割った合計値です。例えば、100人中65人の顧客が5段階評価で4または5と評価した場合、UESは65となります。

6. タスク成功率 (TSR)

👉 ユーザーがタスクをどの程度うまく実行できるかの指標

📋 ユーザーインタビュー (mod/unmod)

🎯 %表示なので、高いほど良い。

🕒 機能リリース時

🔢 Calculator. ユーザーはタスク完了時に採点されます。各機能/タスクのTSRは、タスクを完了または機能を使用したユーザーの割合です。

7. 効率性 (タスクを完了するまでの時間)

👉 ユーザーが特定のタスクを完了するのにかかる時間の目安

📋 ユーザーインタビュー (mod/unmod) 可能であれば監査報告書、またはツールの使用

🎯 時間がかからないほど良い。

🕒 機能リリース時

🔢 これは、ユーザーがタスクを完了するのに費やす時間を観察することで算出されます。

8. データ管理におけるエラー率とデータ損失

👉 ユーザーが機能を使用する際に遭遇するエラーの数、またはシステムが生成するエラーログの数を示す

📋 定義されたエラー状態に応じたユーザーテストと監査ログ

🌟 少ないほど良い

🕒 機能リリース時

📊 エラー率は、ユーザーがエラーに遭遇した回数を合計することで算出されます。

収集されるユーザーデータ：

これは、より幅広い文脈で上記の指標を表示し、指標の評価を行うために必要です。例えば、異なるユーザーロールのシステムユーザビリティスコアを表示する場合などです。

- ユーザーの役割：異なるユーザーの役割におけるユーザー体験指標を分析し、強化が必要なものを特定する。
- フィードバックが製品/プラットフォームのユーザーベースを代表するものであり、1～2の組織による結果ではないことを確認するための組織名。
- アプリケーションへの依存度を文脈化するための使用頻度
- セッション時間、使用頻度と同じ
- デバイスの種類、UXの不具合をデバッグし、より良い体験を提供するためにどのデバイスを優先すべきかを理解する
- ブラウザの種類、デバイス種類と同じ

ユーザビリティフレームワーク

1. ユーザビリティ指標ダッシュボード
 - a. それぞれの調査後に測定値を追跡します
 - b. ダッシュボードでは、指標の経時的な変化のグラフによる傾向を示します。
 - c. すべての指標の計算は個別に表示されます。
2. ユーザビリティフレームワークを導入するためのプロセス
 - a. アンケートで回答を収集する
 - i. 以下の情報を持つユーザーのデータベースを作成する：名前、ユーザーの役割、組織、eメール。
 - ii. アンケート用紙を四半期ごとにeメールで送信する。自動メール送信ツールを使用すると、四半期ごと、または必要に応じて随時、アンケート用紙を簡単に送信することができます。

- iii. すべての回答を集めるために2週間を確保する。アンケートを送信してから1週間経っても回答がないユーザーにはフォローアップする。
- iv. 効果的な指標算出のため、アンケートに対するユーザー回答率を75%と想定する。
- b. 指標の記録と計算
 - i. アンケート実施ごとに、指標計算機を使用してメトリクスを計算し、トラッカーに記録します。
- 3. 推奨事項や提案を導き出すためのユーザビリティ・テストのガイドラインを設定する
 - a. スコアに基づいて、ユーザーが特定の反応を示した理由、ユーザーにとって何が有効で何が有効でないかを調査するための詳細なユーザーテストセッションを実施する
 - b. セッションの所要時間は、機能/タスクの複雑さに応じて15分から45分とする。
 - c. テストセッションを実施する際には、付録Aのテストテンプレートを使用する。
 - d. ユーザーから直接寄せられた回答は、製品の改善に役立てられる。

留意事項

- 1. 評価指標そのものは、ユーザー体験を向上させるための洞察を提供しない。洞察を得るためには、ユーザビリティ・テストを実施する必要がある。
- 2. 評価指標はユーザーの回答に直接基づいているため、ユーザーとのつながりがスコアに影響を与える可能性が高い。
- 3. ユーザーのセキュリティ・スコアは額面通りに受け取ってはならず、技術的なセキュリティ監査と照合すべきである。ユーザーの感覚と安全プロトコルの両方は、アプリケーションのセキュリティ状態を表現する貴重な情報です。
- 4. このフレームワークは、HROが製品/プラットフォームを使用するすべてのユースケースに適用できます。ユーザーは自身の経験に基づいて評価を行い、自身の使用事例に基づいてフィードバックを提供します。そのフィードバックは、製品ビジョンや使用事例によって、製品ロードマップに反映される場合もあれば、されない場合もあります。
 - a. Case: ユーザー満足度調査にユーザーが否定的な反応を示した場合、その使用事例やフィードバックが製品ビジョンと一致していない場合は、その回答を無視するという選択肢もあります。その推論は、そのユーザーの利用ケースが製品/プラットフォームに最適ではないというものであり、製品/プラットフォームのユーザー体験が良くないというものではないかもしれない。
- 5. 製品/プラットフォームのユーザー数が少なすぎる場合、サンプル数が少ないと指標の結論が導き出せない可能性がある。

6. 理想的には、ユーザーテストを行うまでは、機能とテストガイドラインは変更すべきではないすべてのユーザーに対する作業は完了しました。例えば、ユーザーテストの実施中にインストール手順がさらに最適化され、異なるユーザーに対して異なるUXが提供される場合、指標は一貫性のない結果につながります。

枠組みガイドライン

1. 新機能が導入された場合、ユーザビリティ指標ダッシュボードを拡張することができます。これを行うには、まず付録Aに新機能のテストガイドとともに機能を追加し、次にタスク成功率と頻度トラッカーに追加します。
2. 用語「顧客」と「ユーザー」は、互換的に使用されています。

Appendix-A: タスク/機能とテストガイド

Task/Feature	Testing guide (WIP)	Timing for User effort score(UES)	Comments
Task 1	Testing guide link		
Task 2	Testing guide link		
Task 3	Testing guide link		
Task 4	Testing guide link		
Task 5	Testing guide link		
Task 6	Testing guide link		
Task 7	Testing guide link		

Task 8	Testing guide link		
Task 9	Testing guide link		
Task 10	Testing guide link		