

2章 信頼性

～なぜ確信できなくてはいけないか？

ルッツ・プルシェルト, マリアン・ベトリ

Making Software

エビデンスが変えるソフトウェア開発

Andy Oram, Greg Wilson 編

久野禎子, 久野靖 訳

pp.17-33

エビデンス

- 証明ではない

- ある説明が正しくない確率より正しい確率のほうが高そうである

- 別の説明より正しそうであると結論づけさせるに足るような、実践に基づく何らかのデータ

ソフトウェア工学におけるエビデンス

- ソフトウェア工学は科学的調査（エビデンス）と聡明な実践者たちの報告（体験の自己観察）に基づいている
- ただし、自己観察がその他のエビデンスと整合しているかについては熟考の余地がある

→エビデンスをいかに信頼性や目的への適合性に照らして評価するか

信頼性 (Credibility)

- 提示されたエビデンスとそれについての主張をあなたが信じるか（ないし、信じなければならないか）の度合い
- 信頼性の一部であるような概念に妥当性がある
- 妥当性とは、そのエビデンスに基づく調査や主張が、関心対象の現象をどれくらい正確に表現しているかの度合い

妥当性 (Validity)

- 観察したことが、観察しなかったこと、観察していたと思ったものと一致しているかどうか
- 使用した尺度が、測ることを意図したものを実際に測っているかどうか
- 起きたことがなぜ起きたかについての説明が正確かどうか
- 調査した事項を、他の概念的に比較可能な設定へと、一般化できるかどうか

適合性

- あなたがエビデンスとその内容に興味を持つ（ないし、持つ必要がある）度合い
- 例外
 - 調査内容はあなたにとって興味があるのに、その解答が得られた環境はあなたの環境とかなり異なっている場合
 - この場合、適合性はその結果があなたの環境に一般化できる度合いのことを言う

量的エビデンスと質的エビデンス 1/2

- 量的エビデンス
 - 測定が中心
 - 「比較(comparison)」 「もし~なら~(if)」
 - 「どれだけ(how much)」
- 質的エビデンス
 - 「なぜ(why)」 「どうやって(how)」

量的エビデンスと質的エビデンス 2/2

- 量的エビデンス

Merit : 状況を少量の簡潔な記述で補足でき、物事を明快化

Demerit : 結果が本当に意味していることがわかりにくい、その結果が適応可能である場合の判断が難しい

- 質的エビデンス

Merit : 現実世界の複雑さを反映し明らかにする

Demerit : 評価が難しい

エビデンスを組み合わせる

- エビデンスを組み合わせることで、異なる形式のエビデンスや独立した情報源からのエビデンスが合致するなら、それらの「重み」が合わさって、より信頼できるものとなる
- 例：ベンチマーキング
 - 互いに統合できるような結果をもたらすような、ソフトウェア工学手法の一例

エビデンスの種類, その強さと弱さ

- AWE(A Wonderfulnew Excitement)
 - 「素晴らしく, わくわくする物」
- BURP(Boring but Usually Respected Predecessor)
 - 「退屈だけれどおおむね尊重すべき前例」

方法1. 制御実験と擬似実験

- 信頼性
 - 制御実験の信頼性はかなり高く，結果の解釈が極めて明確
 - タスクの偏り
 - グループ編成の偏り
- 適合性
 - 精密であり設定範囲が狭いので，**外の世界に適用しづらい**

方法2. 質問調査

- 信頼性

- 高い信頼性を得るのは困難
- 漠然としていたり，誤解しやすかったり，あいまいな質問
- 調査が複雑な事実に関する問題についての質問

- 適合性

- 最も興味のある質問ができないことがあったり，回答がうまく解釈できず一般化が難しいことがある
- コンテキストを提供することで有益である可能性がある

方法3. 経験報告と事例研究

- 信頼性

- その研究のコンテキストに固有のもの
- 記録されているものの選別性の問題
- あいまいさの問題

- 適合性

- ある環境と別の環境とでは多くの重要な違いがあり
　　そうで適合性の解釈は難しい
- 結論の適合性については確認が難しい
- 質問の適合性については認められることがしばしば
　　ある

報告において信頼性を示すもの 1/2

- 漠然としておらず詳細で，曖昧ではなく正確である
- 議論の進め方が取り繕ったり飾ったりせず正直で，論理と論証が正当である
- 調査の設定が単純でありすぎたり過度に特殊化しておらず最大限の適合性を持たせるようにしている
- 研究報告の書き方として飾り気がなくても魅力がある

報告において信頼性を示すもの 2/2

- 明確なリサーチクエスチョン
- 調査設定の十分な記述
- 意味があり理解が可能なデータ表現
- 限界を正直に議論する

まとめ

- 高い信用性は容易には得られない
- ソフトウェア工学技術者は、メディアや報告書を要約し、信頼性を判断し、自分に適合する結果を持ち帰ることができる
- 様々な形式のエビデンスを受け入れ、それらについて真剣に考えるべき

所感

- これから論文を読んでいく上で、信頼性や適合性について、どれくらい信頼できるか、何に適合できそうかなどについて考えながら読んでいこうと思った