

Problemas metodológicos em Ciência da Computação

Jacques Wainer

Instituto de Computação
UNICAMP

A minha inquietação

Eu não acho que a nossa comunidade sabe fazer ciência de qualidade.

Com a exceção de algumas sub-áreas (teoria, otimização, etc), a pesquisa em Ciência da Computação (CC) se resume a fazer artefatos (programas, sistemas, modelos, metodologia de desenvolvimento, etc) que (talvez) resolvam um problema.

Justificativas em CC

Como se justifica uma pesquisa/linha de pesquisa em CC

- I) este artefato resolve melhor¹ este problema¹
- II) este artefato vai trazer benefícios para pessoas/empresas
- III) este artefato nunca foi feito antes e é promissor
- IV) este artefato resolve este problema²
- V) este artefato vai ser útil para outros desenvolvedores

Justificativas em CC

Algumas dessas justificativas requerem **avaliações**

- I) este artefato resolve melhor¹ este problema¹
- II) este artefato vai trazer benefícios para pessoas/empresas
- III) **este artefato nunca foi feito antes e é promissor**
- IV) este artefato resolve este problema²
- V) este artefato vai ser útil para outros desenvolvedores

(I) Resolve melhor

Ex: algumas áreas de IA, otimização,

Melhor¹ = métrica bem definida, Problema¹ = dados, exemplos.

- comparar o artefato com as alternativas nos dados/exemplos
- medir a métrica
- questões de inferência estatística - intervalos de confiança, p-value, técnicas estatísticas (t-teste, anova, tabelas de contingência, ROC, etc).

(II) Vai trazer benefícios

Ex: educação suportada por computador, informática médica...

Até quando nós vamos poder dizer que vai trazer benefícios e não mostra-los?

Medir benefícios:

- Avaliações qualitativas e quantitativas.
- Desenho experimental
- Técnicas estatísticas de análise dos dados.

Analise qualitativa

Os benefícios não estão bem definidos.

- avaliar o impacto da introdução do sistema na empresa (técnicas etnometodológicas, análise comparativa, técnicas de estudo de caso em administração(?)).
- medir a atitude das pessoas quanto ao sistema (técnicas de criação, aplicação e análise de questionários, entrevistas, focus groups, etc)
- analise correlacional entre resultados de empresas/pessoas que usam e não usam o sistema (técnicas de modelagem estatística)

Analise quantitativa

Benefícios mensuráveis.

Diferença de avaliações em laboratório (*in vitro*) e no ambiente real (*in vivo*).

Toda literatura em ciências sociais sobre desenho de experimentos:

- validade interna e externa
- um grupo versus mais de um grupo
- desenhos experimentais e quase-experimentais
- pré-teste e pós-teste
- efeitos placebo e Hawthorne, experimentos duplo cego.

Outros problemas

Quem faz o sistema deveria avaliá-lo? Isso é um problema sério em ciências médicas.

- Redes de avaliação: um avalia o sistema do outro, mas não mutuamente!
- Avaliar e não implementar nada vale como um mestrado ou doutorado em CC?

(III) Novo e promissor

Ex: IA, groupware, inicio de todas subáreas...

Como prover evidências que isso não foi feito antes?

- Em ciências médicas, mostra-se o resultado de sequências pesquisas em indexadores de publicações (MEDLINE) (com vocabulário controlado)
- Fazer o mesmo no Google Scholar/citeseer/sirus/ACM DL ? Que fazer quanto ao vocabulário não controlado??

Promissor???

(IV) (também) Resolve o problema

Problema² = descrição genérica de um problema *teoricamente* real (mudança dinâmica em workflows, etc).

O artefato novo (também) resolve (a sua versão do) problema.

- Não há uma avaliação - apenas menciona-se as alternativas na seção de trabalhos relacionados.
- Indica talvez áreas ainda imaturas.
- Sofre o problema de “moda”
- Cria uma babel de soluções.

(V) Vai ser útil para outros

Pressupõe que é novo. De novo a questão de “provar” que é novo.

Como medir utilidade?

- Tornar público e contar o número de downloads?
Pedir que usuários mencionem o sistema nos agradecimentos? Cartas de agradecimento de usuários?
- Custo de manter o software através de sua história!!

Isso é um problema brasileiro?

Limitações metodológicas não são um problema de CC em geral e não apenas Brasil?

Talvez. Mas eu vejo uma tendência em direção questões de avaliação, em diferentes graus, em algumas áreas onde eu atuo:

- IA-aprendizado: intervalos de confiança
- CSCW: técnicas de avaliação qualitativas - etnometodologia; descrição do problema e não da solução.
- Informática médica: excelente exemplo. Quase nada é publicado sem uma avaliação, normalmente quantitativa e *in vitro*, tendência é para avaliação *in vivo*.

Mais realista que o rei?

Nas áreas onde a preocupação com avaliação ainda não é forte, podemos iniciar o processo.

Não só adaptar metodologias que já existem, mas criar novas.

Avançar em outras áreas - não parece haver o correspondente ao MIS no Brasil (parte da administração?).

Qual é meu desafio/desejo

Capacitar a comunidade acadêmica em CC em métodos de avaliação. Cursos, livros: para pesquisadores e alunos.

Convencer partes significativas da comunidade a centrar nas questões de avaliação e não de construção. Redefinir parcialmente o que é um mestrado e doutorado.

Construir redes de avaliação.