

Um Grande “Tour” sobre Métodos Empíricos

Cleidson de Souza
cdesouza@ufpa.br

Tópicos Especiais em Engenharia de Software
Departamento de Informática
Universidade Federal do Pará

Agenda

- Introdução
 - Empiricismo
- Métodos de Coleta e Análise de Dados
 - Métodos Quantitativos
 - Definição
 - Exemplos: Experimentos, Surveys, Redes Sociais,
 - Métodos Qualitativos
 - Definição
 - Exemplos: Estudo de Campo e Etnografias;
- Triangulação

2

Empiricismo

- Evidência empírica refere-se a observações que as pessoas experimentam através dos sentidos: tato, visão, audição, paladar e olfato.
- O objetivo é confirmar teorias existentes ou gerar novas teorias.
- Exemplo:
 - Engenharia de software empírica é baseada em evidências (dados) de projetos de desenvolvimento de software reais;
 - Existem revistas e conferências dedicadas ao assunto;

3

Empiricismo

- Teoria da relatividade de Einstein: energia = massa x c^2 (c =velocidade da luz)
 - Descrição matemática; e
 - Validação empírica; (vários e vários anos depois)
- A evidência empírica coletada visava validar a teoria de Einstein;
- Também pode-se usar evidências empíricas para gerar novas teorias.

4

Métodos para Coleta e Análise de Dados

- Duas abordagens principais:
 - Métodos quantitativos
 - Quantidade;
 - Quantas vezes isto ocorre?
 - Com que frequência?
 - Os dados coletados se encontram na forma de números;
 - Métodos qualitativos
 - Qualidade;
 - Que tipo de coisas ocorrem?
 - Os dados coletados encontram-se na forma de palavras e / ou figuras

5

Métodos para Coleta e Análise de Dados

- Quando usar cada abordagem?
 - Métodos quantitativos, quando você sabe exatamente o que perguntar, o fenômeno que você quer estudar
 - Número de mensagens de texto através do celular enviadas entre jovens entre 12 e 18 anos;
 - Métodos qualitativos, quando você quer saber o que perguntar
 - Qual o objetivo de enviar mensagens pelo celular?

6

Métodos Quantitativos

- Adotada por pesquisadores positivistas
 - Baseada na mesma abordagem e tradição das ciências naturais como física e química;
 - favorecem a pesquisa “objetiva”, tentando medir precisamente aspectos das pessoas e testar hipóteses através da análise cuidadosa dos números obtidos a partir da coleta de dados.
 - O objetivo é encontrar modelos ou padrões que descrevem um determinado fenômeno; a lei da gravidade, leis da manutenção de software, etc.

7

Métodos Quantitativos

- Importante fundamentação matemática é necessária para a análise dos dados
 - **Diversos** testes estatísticos podem ser utilizados para avaliar as hipóteses de interesse;
 - Chi-quadrado, A-nova, t-test, etc etc etc etc.
 - Estatisticamente significante: a probabilidade que o resultado de um estudo seja devido ao acaso;
 - Geralmente adota-se um limite de .05 ($p < .05$), ou seja, a probabilidade de rejeitar a hipótese de interesse é de 1 em 20;
 - Quanto maior a probabilidade de rejeitar a hipótese de interesse, maiores são as chances do resultado não se dar por acaso;

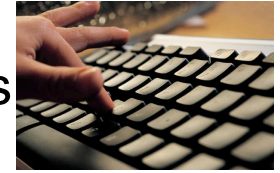
8

Coleta de Dados: Métodos Quantitativos

- Experimentos
 - Qual a interface com usuário mais apropriada para pessoas aprendendo a utilizar a internet? Internet Explorer ou Firefox;
 - Qual a técnica mais efetiva para inspeção de software: leitura baseada em checklist ou leitura baseada em xxxx?
- "Surveys"
 - O papel de mecanismos formais e informais de coordenação no sucesso de projetos de desenvolvimento de software;
 - Aspectos culturais que influenciam autores de blogs;
- Análise de Conteúdo (Content Analysis)
 - Técnica para examinar informações ou o conteúdo escrito ou simbólico de um determinado material;
 - Contar o número de vezes que uma certa palavra ocorre em conversas de chat entre adolescentes.
- Estatísticas existentes (Existing Statistics)

9

Experimentos



- A produtividade dos programadores aumenta com a utilização de dois monitores?
- Projeto:
 - 10 programadores usando 1 monitor
 - 10 programadores usando 2 monitores
 - Programar um problema de dificuldade média em 1 hora
 - Análise dos programas ao final do prazo
- Controle: habilidade de programação, IDE, tarefa a ser executada, produtividade (qualidade do código / número de erros, etc etc etc.

10

Survey



- Utilização de mensagens instantâneas no celular
- Questionário:
 - Quantas mensagens você manda por dia?
 - Quantas mensagens você envia por dia?
 - Quantas pessoas você telefona por dia?
 - Quantas pessoas te telefonam por dia?
 - Quantas pessoas você tem cadastradas na agenda do seu celular?
- Análise:
 - Pessoas que fazem mais ligações mandam mais mensagens?
 - Pessoas com uma agenda maior mandam mais mensagens?

11

Survey - 2

- Impacto de mudanças nas atividades de programação
 - Quando um trecho de código de seu colega muda, qual a implicação desta mudança no seu próprio código?
 - Gasto pouco, algum, ou muito tempo fazendo modificações no meu código?
 - O esforço para fazer estas modificações é? Pequeno, médio ou grande;
- Análise
 - Mudanças de colegas na minha sala me afetam menos, do que mudanças de colegas distribuídos;

12

Redes Sociais

- Uma rede social consiste de um conjunto finito de atores e a(s) relação(ões) definidas entre eles (Wasserman and Faust 1994, p. 20).
- Exemplos de relações incluem laços familiares, amizades, contextos de trabalho, confiança, dependência, etc.

13

Redes Sociais (2)

- São representadas graficamente por sociogramas. Um sociograma é um grafo com nodos e arestas.
 - Nodos indicam atores sociais enquanto que arestas representam os relacionamentos entre os atores.
 - Atores normalmente são pessoas, mas também podem ser times, departamentos, organizações inteiras, etc.

14

Redes Sociais - Exemplos

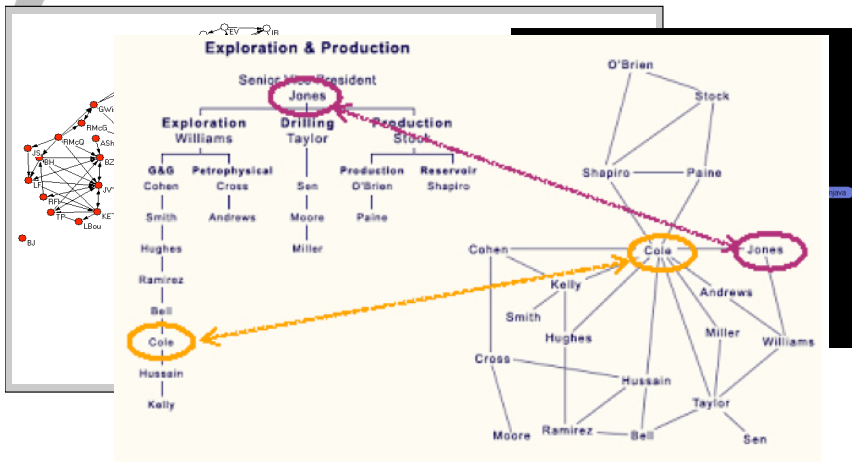


Figure 2-2. ReferralWeb.

15

Redes Sociais - Utilização

DOMÍNIOS

- Propagação de doenças;
- Combate ao terrorismo;
 - Razão do aumento das pesquisas nos últimos anos depois do "11 de Setembro";
- Difusão de tecnologias;
- Análise organizacional;
 - Inovação, desempenho, criação de comunidades, escândalo da Enron nos EUA,
- Estudo de comunidades virtuais;
- Etc.

APLICAÇÕES

- Identificação de "experts" na organização
 - Quem você procura quando precisa de ajuda técnica?
- Oportunidades para Colaboração
 - De quem você depende para realizar seu trabalho?
 - Engenheiros de software em uma organização tinham as mesmas dependências mas não estavam colaborando;
- Pessoas que conectam grupos separados, atuam como "bridges";

16

Redes Sociais - Coleta de Dados

- Em um projeto, pode-se perguntar:
 - Com que frequência cada par de integrantes de um time tem conversas informais?
 - Todo dia, uma vez por semana, uma vez por mês, etc.
 - Quem confia em quem?
 - Quem depende do trabalho de quem?
 - Quem está ciente do trabalho de quem?
 - Quem é amigo de quem?
 - Etc, etc.
- Também é possível extrair as redes a partir de *logs* de ferramentas

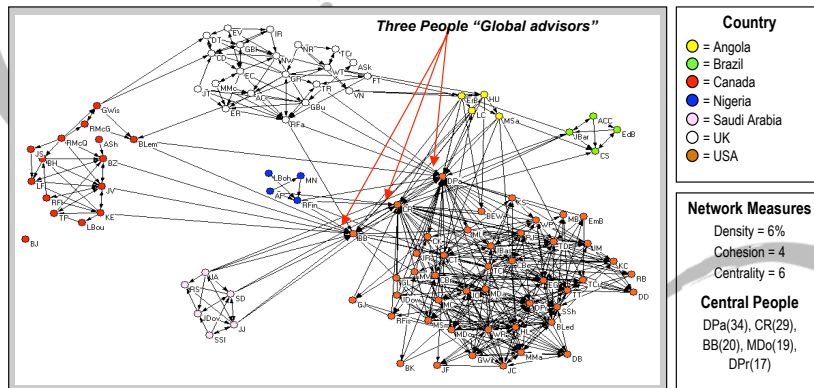
17

Análise de Redes Sociais (SNA, social network analysis)

- Diversas métricas podem ser calculadas em uma rede social
 - Densidade da rede
 - # de conexões existentes / # possível de conexões;
 - Grau (de entrada ou saída) de cada nodo;
 - Atores que recebem várias conexões ou que tem várias conexões;
 - Centralidade de cada nodo;
 - Vários tipos de centralidade..
 - Degree, closeness, betweenness, flow, etc.
 - *Hubs / bridges* ou “pontes”;
- Observação: Matrizes são utilizadas ao invés de grafos para a manipulação dos dados.

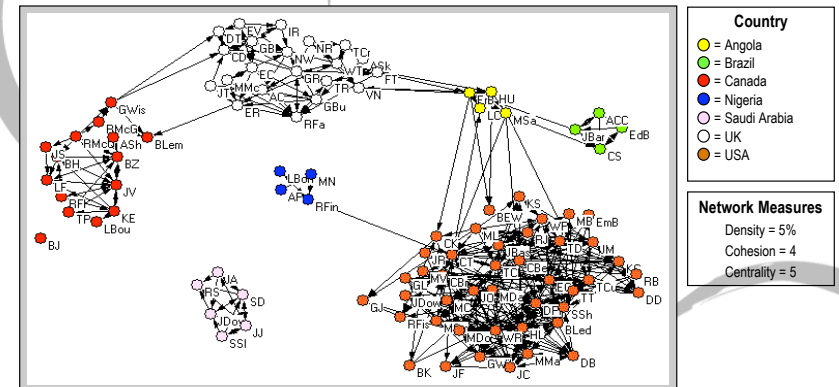
18

Redes Sociais: Desenvolvimento Distribuído de Software



19

Redes Sociais: Desenvolvimento Distribuído de Software



20

Métodos Quantitativos

- Exemplos de artigos:
 - Experimentos
 - Thelin, T.; Runeson, P.; Wohlin, C.; Olsson, T.; Andersson, C. *How much information is needed for usage-based reading? A series of experiments*. Proceedings of the Empirical Symposium in Software engineering, pages: 127- 138, 2002.
 - Surveys
 - Kraut, R. and Streeter, L. *Coordination in software development*. In Communications of the ACM, volume 38(3), pages: 69-81, Mar. 1995.
 - Redes Sociais
- Conferências:
 - ISESE - International Symposium in Empirical Software Engineering
 - CHI - computer human interaction

21

Métodos Qualitativos

- Adotada por pesquisadores interpretativos (interpretive researchers):
 - Ação não tem significado que possa ser diretamente observado. Ela adquire significado entre as pessoas que compartilham o mesmo conhecimento sobre o domínio, o que permite que eles interpretem a ação como algo relevante naquele contexto social;
 - O mundo **social** não está “lá fora” esperando pra ser descoberto independentemente da consciência dos atores, como nas abordagens das ciências naturais;
 - “o mundo é integrado empiricamente ... queira o pesquisador goste ou não” [Glaser 1998, p. 189].

22

Métodos Qualitativos

- Análise dos dados é tão rigorosa quanto nos métodos quantitativos, mas feita de maneira diferente;
- Geralmente requer um período mais longo de coleta e análise de dados;
- Exemplo de abordagem: teoria fundamentada em dados (grounded theory);

23

Métodos Qualitativos: teoria fundamentada em dados

...	
Q: Do you mean that you can describe [architecture] better with PowerPoint?	
A: You can do it much better with it. You can draw empty boxes with Rational Rose, but it isn't as visual.	Problem: tool constraints
Q: Is it a problem of looking good?	
A: When you are presenting it to salesmen and customers, yes it is.	Problem: visual appearance Stakeholder: customer management & marketing Stakeholder: customer
Q: Is it important that you show pretty pictures to customers?	
A: Yes. Especially when you can tell with that picture what you have been thinking. In addition, many times when we are making requirements documents, the customer wants architecture documentation as a PowerPoint presentation. They present the architecture also to other possible suppliers and they do not want to redraw the pictures. It is little like giving a tool to the customer too.	Problem: communicating meanings Rationale: communicating Stakeholder: customer Stakeholder: other suppliers
...	

Figure 2. An example of open coding

24

Métodos Qualitativos

- Exemplos:
 - Pesquisa de campo (Field Research)
 - Estudos de casos que normalmente ocorrem durante um período longo de tempo;
 - Pesquisa histórico-comparativa (Historical-Comparative Research)
 - Examina aspectos da vida social em um período histórico ou através de diferentes culturas.
 - Pesquisadores podem focar em um ou mais períodos, comparar uma ou mais culturas, ou combinar períodos e culturas;

25

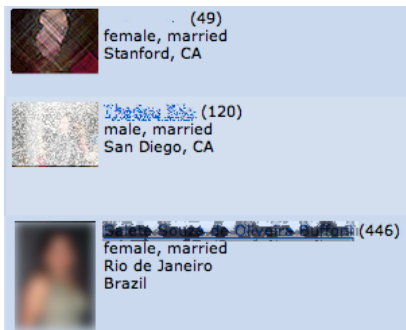
Métodos Qualitativos

- Técnicas de coleta de dados geralmente utilizadas com métodos qualitativos:
 - Entrevistas não-estruturadas ou semi-estruturadas;
 - “Filmagem” dos informantes;
 - Notas de campo (field notes) baseadas em observação participativa ou não-participativa (em inglês também é conhecido como “shadowing”); e
 - Coleta de documentos.

26

Estudo de Campo

- Como as pessoas utilizam o orkut no seu dia a dia?
 - Para se comunicar com os amigos atuais?
 - Fazer novos amigos?
 - Para achar antigos amigos?
 - É sinal de status ter muitos amigos no orkut?
 - **O que significa estar no orkut?**
- Entrevistar e observar as pessoas, no dia a dia, utilizando o Orkut.



27

Etnografias

- Antropologia
 - Estudos em lugares não-familiares para descobrir o que se passa;
 - Bronislaw Malinowski com o livro “Argonauts of the Western Pacific”;
 - Índios na Amazônia;
 - Geralmente requer períodos longos de coleta de dados;
- Grego:
 - *ethnos* = nação, pessoas; *graphein* = escrita
 - A escrita de uma cultura
- Não somente o que as pessoas fazem, mas o que elas experimentam, o ponto de vista delas, a motivação, o porquê;

28

Etnografias

- Nos permite obter um entendimento mais rico sobre:
 - Como as pessoas realizam seu trabalho;
 - Como as pessoas trabalham, aprendem, se comunicam, usam ferramentas, se movem, etc;
 - Organização sócio-cultural e realizações do dia a dia;
- **E porquê nós (da computação) deveríamos nos importar?**
 - Para projetar, desenvolver (e vender) soluções que são úteis (de valor) para as pessoas que vão utilizá-las
 - Etnografia industrial: IBM, Microsoft, Intel, (Xerox) PARC, etc.

29

Exemplo de Etnografia Industrial

Intel - Genevieve Bell: utilização de tecnologias de informação (celulares e outros) no dia a dia.



Livros e revistas chineses com instruções de como paquerar usando mensagens instantâneas e como contar piadas. ³⁰

Métodos Qualitativos

- Exemplos de artigos:
 - Rebecca E. Grinter, Using a configuration management tool to coordinate software development, Proceedings of conference on Organizational computing systems, p.168-177, August 13-16, 1995, Milpitas, California, United States.
 - Smolander, K. Four Metaphors of Architecture in Software Organizations: Finding Out the Meaning of Architecture in Practice, Proceedings of the Empirical Symposium in Software engineering, pages: 211- 221, 2002.
- Conferências:
 - CSCW - computer supporter cooperative work
 - ECSCW - european CSCW

31

Qual abordagem utilizar?

- Não existe método mais correto ou mais apropriado, tudo depende do problema a ser resolvido e do contexto.
- A utilização de um método não exclui a utilização de outro. Pode-se usar mais de um método para coleta de dados, o que é chamado de triangulação.

32

Que abordagem utilizar?

- Triangulação:
 - Utilizar diversos métodos de coleta de dados para averiguar a autenticidade dos resultados;
 - Adotar uma variedade de abordagens na pesquisa de um problema, como se você enxergasse o mesmo problema de diversos ângulos. Através destes diferentes métodos, diferentes resultados serão obtidos. Estes resultados deverão ser similares e formar um triângulo. Dentro deste triângulo, a “verdade” se encontra.

33

Observação Importante!

- A distinção entre dados quantitativos e qualitativos deve ser encarada com flexibilidade:
 - Redes sociais podem ser usadas apenas metaforicamente para a coleta de dados qualitativos;
 - Dados qualitativos podem ser quantificados: contar a ocorrência de uma palavra em uma entrevista;
 - Etc, etc.

34

Projeto da Disciplina

- Times deverão escolher uma abordagem:
 - Surveys ou questionários
 - Elaboração de perguntas;
 - Análise estatística;
 - Redes Sociais
 - Elaboração de perguntas;
 - Análise estatística;
 - Mini-etnografias
 - Requer a condução de entrevistas;
 - Observação;
 - Observação: experimentos não serão considerados porque o tempo da disciplina é curto.

Não subestimem o tempo e esforço necessários para o desenvolvimento deste projeto!

35

Projeto da Disciplina

- Domínio do problema
 - De preferência desenvolvimento de software
 - SERPRO,
 - Polícia Militar,
 - Etc.
 - Mas, opcionalmente, qualquer tecnologia interativa e/ou colaborativa
 - Celulares, sites, PDAs, etc.

Bolsas de IC?

36

Importante!!!!

- Todos devem se registrar em
 - <http://groups.google.com/group/topicos-es-2006>
 - Link disponível a partir de <http://www2.ufpa.br/cdesouza/teaching.html>
- Envio de mensagens:
 - topicos-es-2006@googlegroups.com