



GRUPO DE PESQUISA APLICADA EM
ENGENHARIA DE SOFTWARE

Estudo de Mapeamento Sistemático em Engenharia de Software

Jorge Dias

Abril 2011



DEPARTAMENTO DE CIÊNCIAS EXATAS
CENTRO DE CIÊNCIAS APLICADAS E EDUCAÇÃO
UNIVERSIDADE FEDERAL DA PARAÍBA

Sobre o DCE

- Departamento de Ciências Exatas – CCAE (Campus IV)
 - Licenciatura em Matemática
 - Licenciatura em Ciência da Computação
 - Bacharelado em Sistemas de Informação
- ARiSE– **A**pplied **R**esearch **i**n **S**oftware **E**ngineering
 - Engenharia de requisitos, Linhas de produto, ES orientado a serviços, Arquitetura de Software de Larga Escala, Gerência de serviço, etc.
 - arise.dce.ufpb.br

Pesquisa científica em Engenharia de Software

- Digamos que...
 - Você é um pesquisador na área de engenharia de software
 - Você quer definir uma nova abordagem ou método orientado a objetos

O que você faz?

Pesquisa científica em Engenharia de Software

- Procurar **trabalhos científicos** primários, secundários, etc...
- Estudos secundários
 - **Estudo de mapeamento sistemático**
 - Mais geral
 - Revisão sistemática da literatura
 - Mais detalhado
 - Mais comparativo

Mapping Study

- Na pesquisa científica, o primeiro passo é procurar trabalhos anteriores
 - Paradigma baseado em evidências
 - Deu certo na medicina
- Falta de pesquisa empírica na engenharia de software
 - Em 2004, surge alguns guidelines para aplicar o paradigma baseado em evidência em ES
 - Primeiro mapping lançado em 2007
- Precisa ser “repetível”
 - Por isso precisa ser sistemático

Definição

“um mapa sistemático é um método definido de construir um esquema de classificação e estrutura em um campo de interesse” (PETERSEN et al., 2008).

Porque fazer um mapping study?

- Mapear uma área, evidentemente
 - Identificar principais papers e seus tipos
 - Identificar “gaps”
- Identificar os principais meios de publicação de uma área
- O próprio resultado do mapping é uma contribuição
- Primeiro passo para um aluno de doutorado

O que precisa para fazer um mapping?

- Um tema a ser mapeado
- Perguntas a serem respondidas (*research questions*)
- Um protocolo para guiar os participantes
- O número ideal de participantes são 3 pessoas
 - Executar com uma pessoa apenas diminui a qualidade do mapping

Criando um protocolo para mapping study

- Introdução
- Qual a área de conhecimento que você quer mapear?
- Qual a justificativa para realizar este estudo?
- Identificar **questões de pesquisa** (*Research Questions*)
- Exemplos:
 - Como estão distribuídas as publicações ao longo dos anos?
 - Podemos distribuir as publicações seguindo uma taxonomia? Se sim, como está esta distribuição?

Processo de busca

- Definir como será o **processo de busca**
 - **Busca automática**
 - Definir quais engines de busca serão utilizados
 - IEEEExplore, Google Scholar, Scopus, ScienceDirect, ACM Digital Library, Engineering Village, etc
 - Definir as strings de busca
 - **Busca manual**
 - Definir que conferências, journals, etc
 - Principais conferências/journals da área

Busca Automática

- Os principais engenhos de busca

ACM Digital Library	http://portal.acm.org/
IEEE Xplore	http://ieeexplore.ieee.org/
Scopus	http://www.scopus.com/
ScienceDirect	http://www.sciencedirect.com/
Engineering Village	http://www.engineeringvillage.com/

Busca Automática

- Definindo as strings de busca
 - Baseado nas **palavras-chaves** das questões de pesquisa
 - Strings podem ter **sinônimos**, variar em número e em idioma
 - Exemplo: a palavra-chave “Programação Orientada a objetos”

*“programação orientada a objetos”, “POO”, “OO”,
“orientação a objetos”, “object-oriented”, “object
orientation”*

Busca Automática

- Definindo as strings de busca
 - Devemos usar **operadores booleanos** nas buscas
 - OR para sinônimos
 - AND para integrar palavras-chaves
 - Exemplo:

Queremos pesquisar sobre métodos de design em POO.

*(“design” **OR** “método”)*

AND

*(“programação orientada a objetos” **OR** “POO” **OR** “OO”
OR “orientação a objetos”)*

Busca Automática

- Definindo as strings de busca
 - ATENÇÃO: para cada engenho de busca tem uma forma de se criar a expressão
 - Deve-se estudar como criar a expressão booleana em cada engenho de busca

Busca Automática

- Exemplo

- IEEEExplore

- ("design" OR "metodo") AND ("OO" OR "orientação a objetos" OR "programação orientada a objetos")

- Scopus

- TITLE-ABS-KEY(("design" OR "metodo") AND ("OO" OR "orientação a objetos" OR "programação orientada a objetos"))

Definir critérios de inclusão e exclusão

- Define quais os **critérios** para **incluir** ou **excluir** um trabalho do mapping
- Precisam estar bem definidos para não haver **discordância nas avaliações** dos pesquisadores
- Mesmo assim, vai haver! 😊

Busca Manual

- Selecionar apenas as principais conferências sobre o tema
 - É uma busca complementar
 - Se o paper for relevante ele será indexado pelos engenheiros e recuperado na busca automática

Mapear

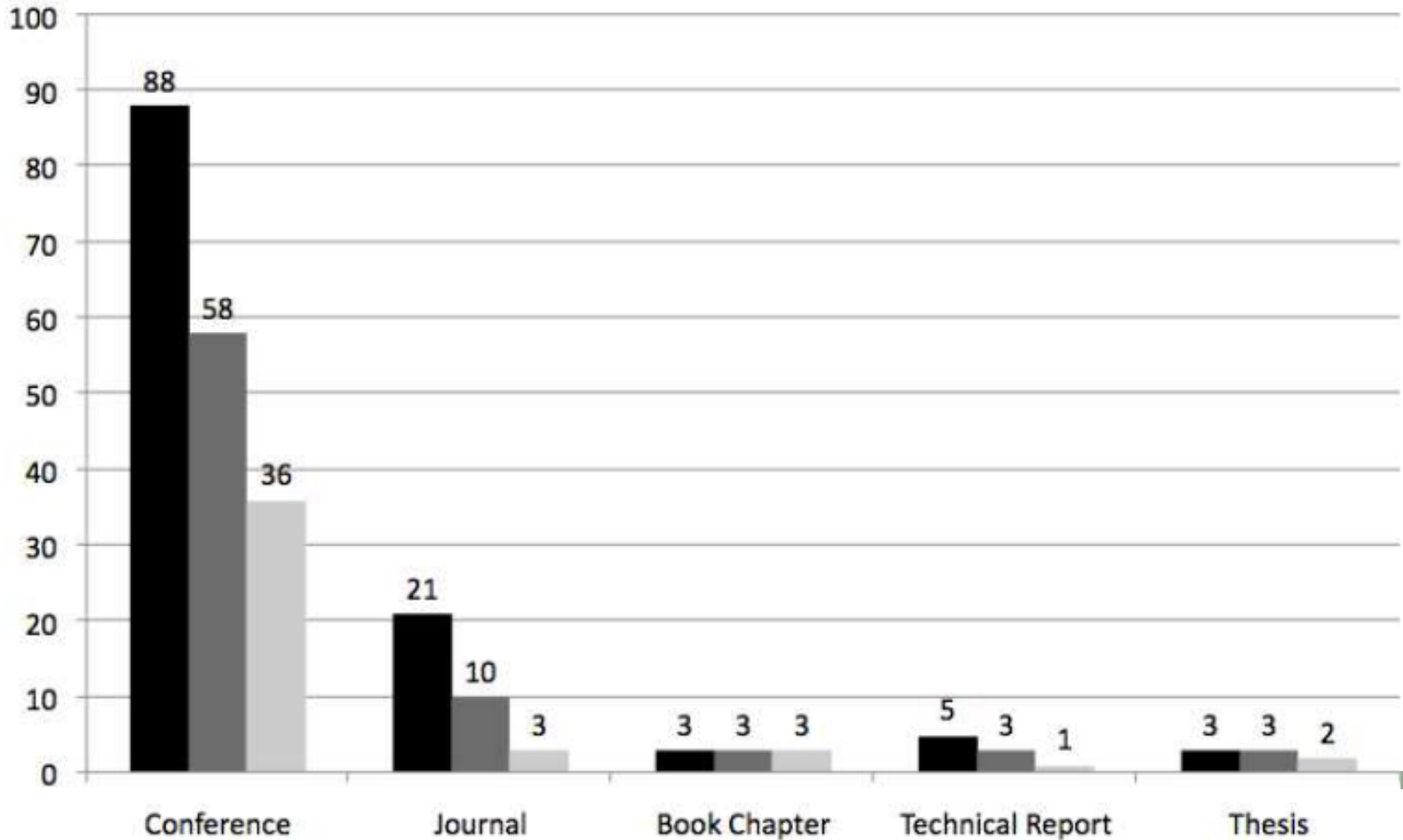
- Analisar os trabalhos selecionados de acordo com as questões de pesquisa
- Criar **taxonomias** para classificar os trabalhos e gerar resultados interessantes
 - Criar uma taxonomia nova não é trivial
 - Tentar aproveitar taxonomias existentes na literatura
 - Utilizar taxonomias das conferências da área
 - Pode usar o DBLP
 - Keywording [Bailey et al. 2007]
- Gerar resultados
 - Geralmente em forma de gráficos

Taxonomias

■ Facetas

Category	Description
Validation Research	Techniques investigated are novel and have not yet been implemented in practice. Techniques used are for example experiments, i.e., work done in the lab.
Evaluation Research	Techniques are implemented in practice and an evaluation of the technique is conducted. That means, it is shown how the technique is implemented in practice (solution implementation) and what are the consequences of the implementation in terms of benefits and drawbacks (implementation evaluation). This also includes to identify problems in industry.
Solution Proposal	A solution for a problem is proposed, the solution can be either novel or a significant extension of an existing technique. The potential benefits and the applicability of the solution is shown by a small example or a good line of argumentation.
Philosophical Papers	These papers sketch a new way of looking at existing things by structuring the field in form of a taxonomy or conceptual framework.
Opinion Papers	These papers express the personal opinion of somebody whether a certain technique is good or bad, or how things should be done. They do not rely on related work and research methodologies.
Experience Papers	Experience papers explain on what and how something has been done in practice. It has to be the personal experience of the author.

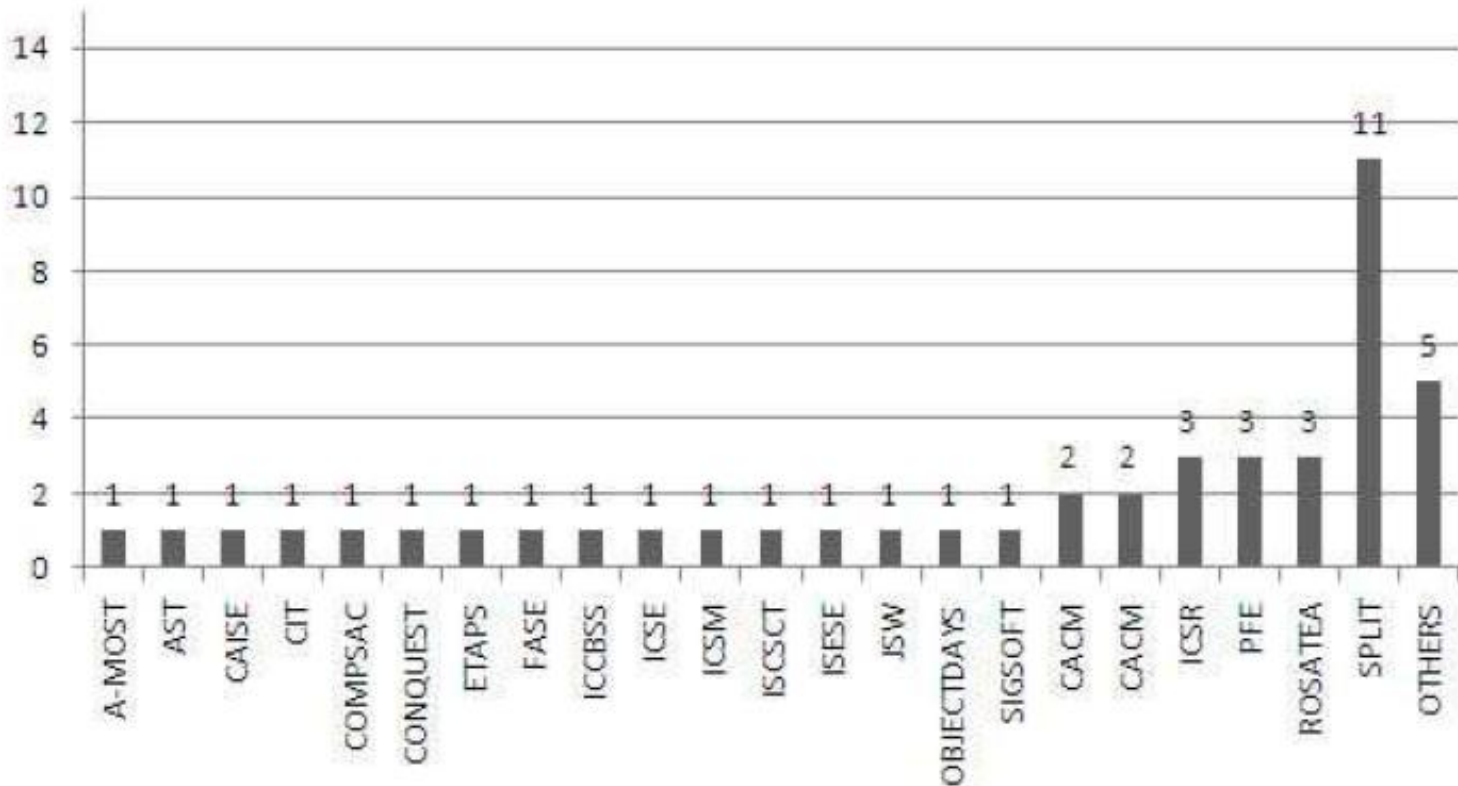
Alguns tipos de resultados



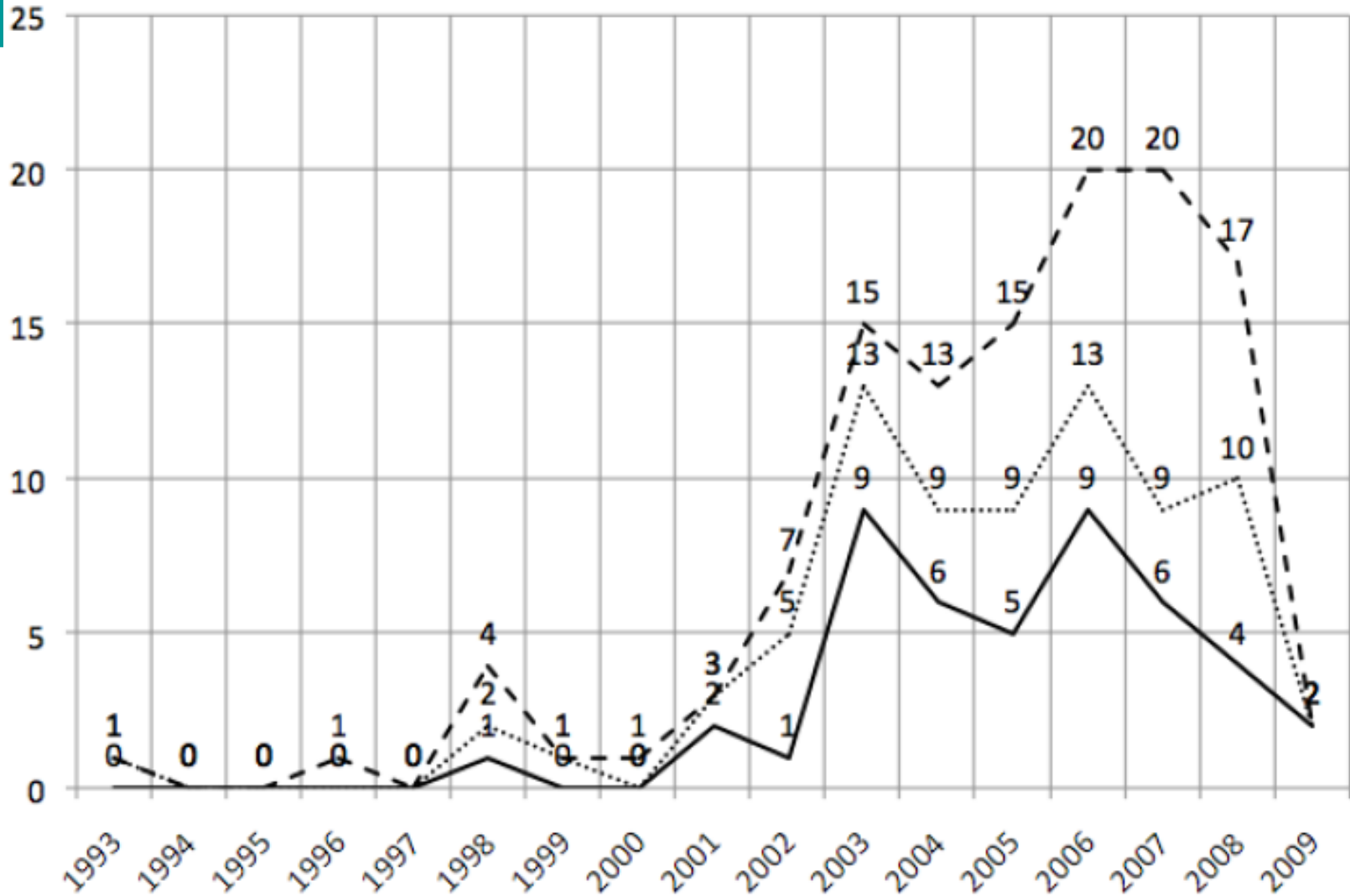
Alguns tipos de resultados

- Estudos por fontes

Studies vs. Sources

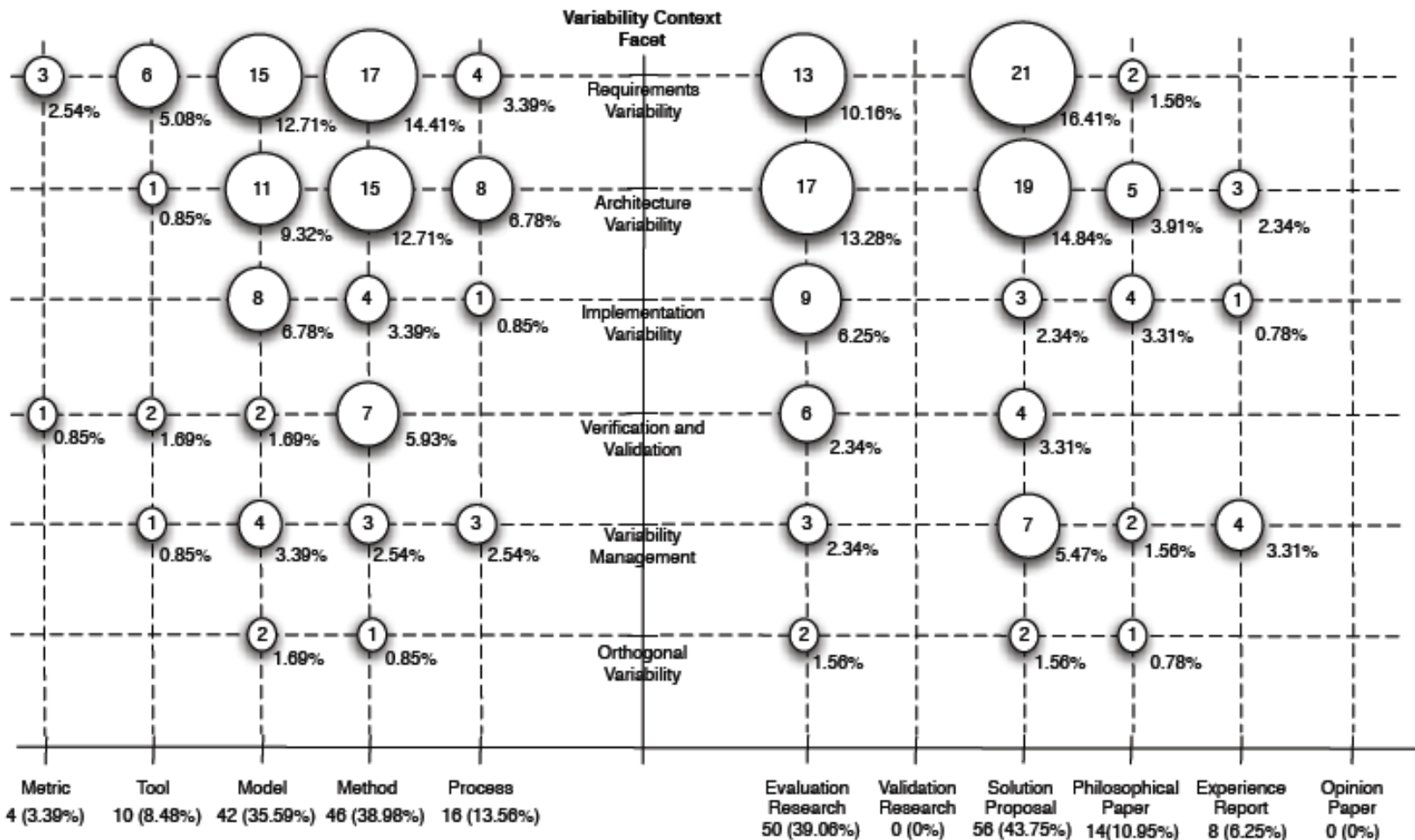


Alguns tipos de resultados



Alguns tipos de resultados

Buble Plot



FAQ

- Devemos ler os papers selecionados por completo?
 - Devemos incluir a “gray literature”?
 - Posso fazer sozinho?
-



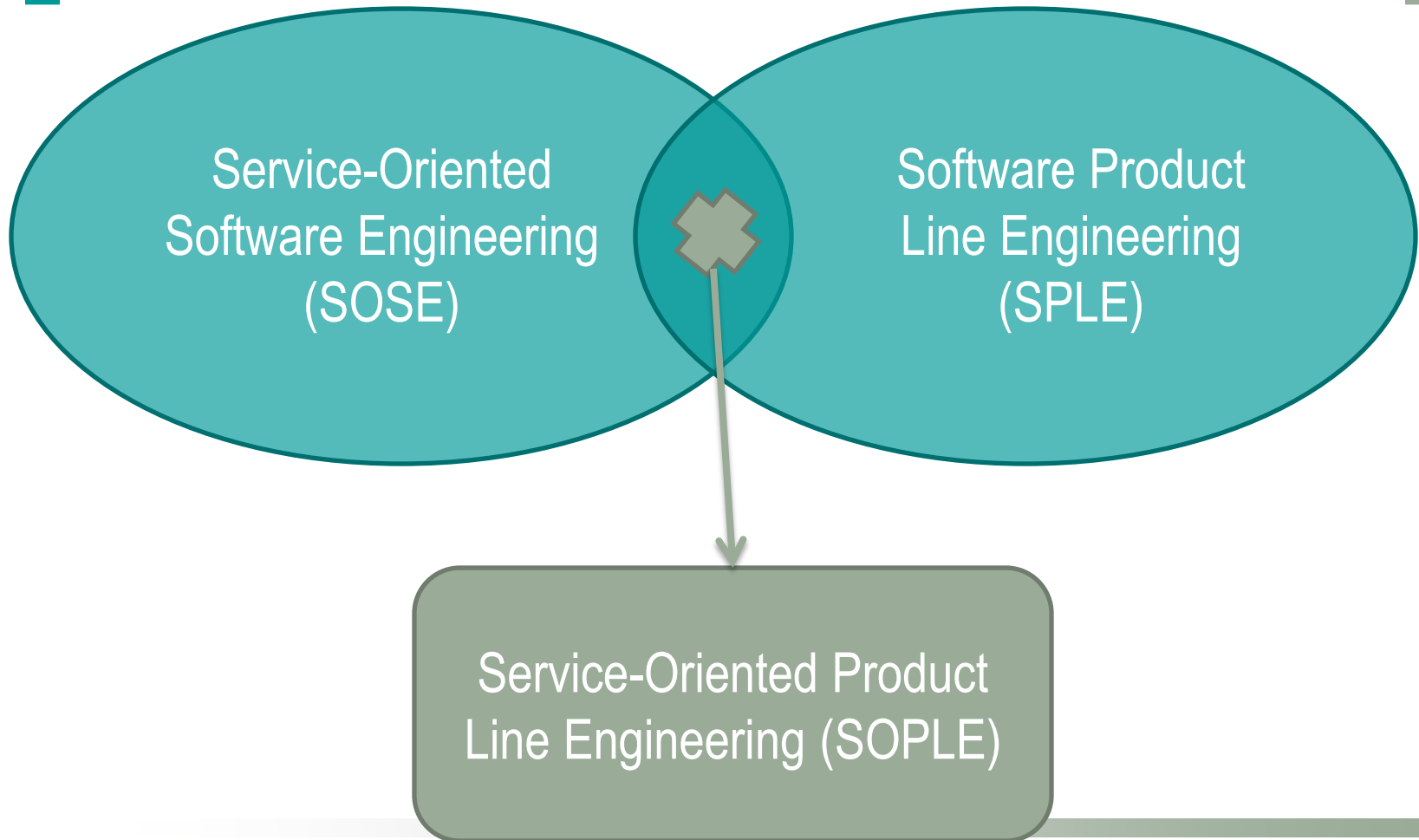
Estudo de caso



Contexto

- Um pesquisador quer encontrar um *gap* para definir seu trabalho de doutorado
- Ele quer investigar a integração de duas áreas da engenharia de software:
 - Engenharia de Software Orientada a Serviços e
 - Linhas de produto de software

Contexto



Research Questions

- **RQ1.** Qual a intensidade de publicações na área de SOPLÉ ao longo dos anos?
- **RQ2.** Quais os objetivos para se integrar as duas abordagens?
- **RQ3.** Que métodos de pesquisa estão sendo utilizados para validar as propostas?
- **RQ4.** Em que domínios e contextos estas propostas estão sendo utilizados?
- **RQ5.** Que tópicos de SOA e SPL estão sendo influenciados nas propostas?
- **RQ6.** Que tipos de variabilidade estão sendo utilizados?

Critérios de inclusão e exclusão

- Critério de inclusão
 - Trabalho que propõe a integração de conceitos de SOSE e SPL. Todos os estudos publicados em journal e conferência que tenha uma proposta SOPLE.

Critérios de inclusão e exclusão

- Critérios de exclusão

- Estudos que não consideram um serviço como sendo uma caixa preta, auto-contido, stateless, fracamente acoplado, interoperável com contratos bem definidos (definição de algum autor renomado da área)
- Pesquisas descrevendo apenas aspectos da área de SOSE sem mencionar linhas de produto
- Pesquisas descrevendo apenas aspectos de SPL sem mencionar SOSE
- Estudos que referem-se apenas a tecnologia web services sem considerar os conceitos de SOSE
- Estudos duplicados. Quando um estudo tiver sido publicado em mais de uma conferência/journal, a versão mais completa será utilizada, isto é, aquela que explicar em um nível maior de detalhe

Alguns Resultados

- Resultados da busca automática sem aplicar critérios

ACM Digital Library	449
IEEE Xplore	282
Scopus	1153
ScienceDirect	264
Engineering Village	444
Total	2592

Alguns Resultados

Etapa	Qtd de Papers
Pesquisa automática sem aplicar critérios	2592
Depois de aplicar o critério de inclusão no título e resumo e excluindo trabalhos duplicados na busca automática	94
Busca manual	58
Interseção entre busca automática e manual	114
Depois de fazer uma análise mais apurada, aplicando critérios de exclusão	X

Principais dificuldades e Lições aprendidas

- O protocolo precisa ser calibrado e sempre atualizado
 - Research questions, Strings de busca, etc...
- Escolher a principal conferência e principais trabalhos da área para ajudar (quando possível)
- É importante gerar resultados parciais
 - Ajuda a identificar novas questões e eliminar outras

Principais dificuldades e Lições aprendidas

- Os abstracts tem sido bastante criticados pela academia
 - *structured abstracts* tenta resolver isso
- O ideal é realizar o mapping com 3 pessoas
 - 1 sempre resolve os conflitos
- A busca da ACM não funcionou no firefox e IE
 - Não me pergunte o porquê 😊

Algumas discussões

- Devemos incluir “gray literature”?
 - Relatórios técnicos, livros, white papers, etc...