

**DON'T
PANIC**

THE HITCHHIKER'S GUIDE TO THE PhD

Andrian (Andi) Marcus

seers group

www.andrianmarcus.net





1995

B.S. Computer Science

1997

MS European Studies

2005-06

Visiting Prof.

**1997-1998
Fulbright Scholar**



THE UNIVERSITY OF MEMPHIS



2001
MS Computer Science

2003
PhD Computer Science





2003-2014
Assistant/Associate Professor

2014-2023
Associate/Fool Professor



2023 – GEORGE MASON UNIVERSITY



WHAT IS MY WORK ABOUT?

"Focus on the students, since graduating great students means you'll produce great research, while focusing on the research may or may not produce great students."

Nico Habermann (1932-1993)

BEFORE WE BEGIN

These are my opinions – albeit not all original

Others may (strongly) disagree – **listen to them!**

There are many ways to do research and often it is highly personal – **remember what Peggy talked about**

Feel free to interrupt and ask questions or comment

WHY ATTEND A SUMMER SCHOOL?

SIESTA 2023

3rd International Software Engineering Summer School

September 11–13, 2023 – Lugano, Switzerland





VISSOFT 2005
Budapest, Hungary





VISSOFT 2005

Budapest, Hungary

MY FIRST SUMMER SCHOOL

7 INTERNATIONAL SUMMER SCHOOL ON SOFTWARE ENGINEERING (ISSSE 2010)

5 - 9 JULY 2010

COMPUTER SCIENCE BUILDING OF THE UNIVERSITY OF SALERNO, VIA PONTE DON
MELILLO, FISCIANO (SA), ITALY

ORGANIZED BY SESA LAB





ICSM 2010
Timisoara, Romania



Context Refactoring Software Systems

- Refactoring: changing software without modifying its external behavior
- Improving non-functional attributes of the software
- Software evolution ... continuous changes
- Changes cause a drift of the original design reducing its quality
- Class Cohesion: how strongly related the various responsibilities of a class are
- Programmers often add wrong responsibilities to a class => its cohesion decreases

EXTRACT Splitting a class with many **CLASS** responsibilities into different classes

REFACTORING

Game Theory Background

Game Theory's captures behavior in strategic situations, in which an individual's success when making choices depends on the choices of others

- A game consists of
 - a set of players (2 or more)
 - a set of moves available to those players
 - payoffs for each combination of moves

The Prisoner's Dilemma

Sally and Tom are accused of fraudulent activities and both want to minimize the time spent in jail

The outcome of this game is represented by the Nash equilibrium (conflict, conflict)

	Tom	Case
Sally	Cooperate	Cooperate
Sally	Defect	Defect
Tom	Cooperate	Cooperate
Tom	Defect	Defect

Payoff matrix for the Prisoner's Dilemma

Game Theory Meets Software Engineering

Modeling Extract Class Refactoring as a Non-cooperative Game

2 PLAYERS Each player is in charge to build a new class refactoring method from the original class

MOVES Iterative game: at each iteration, a player selects a move (one method of the class to be refactored)

PAYOFF Each player selects the method considering the impact on the cohesion and coupling of his class

The game starts by assigning to each player a set of methods having the lowest complexity, i.e., the smallest number of lines of code. The volatility between two methods, i.e., the number of references of methods and variables, is considered as a measure of coupling. The number of methods that a player has to refactor is considered as a measure of cohesion. The game ends when the number of methods that a player has to refactor is zero (the player has refactored all the methods of the class).

Preliminary Evaluation

Case Study Design

Case	System
RQ1	Comparison with AgriMk, JUnitMk
RQ2	Comparison with AgriMk, JUnitMk
RQ3	Comparison with AgriMk, JUnitMk
RQ4	Comparison with AgriMk, JUnitMk

Experiment execution

The refactoring planning is inspired by the iterative refactoring process. We randomly select two classes of one of the subjects and merge them in a single class. We then use the experimental approaches to split the merged class in two classes.

Results (i-Measure)

System	Player	Max/Min	Avg/Std
AgriMk	AgriMk	90%	88%
JUnitMk	JUnitMk	77%	75%
JUnitMk	AgriMk	82%	76%

26th IEEE International Conference on Software Maintenance, Timisoara, Romania, September 15, 2010

ICSM 2010
Timisoara, Romania

ASE 2010
Antwerp, Belgium

A photograph of a man and a woman sitting at a table in front of a brick wall. The man, on the left, has a goatee and is wearing a light yellow button-down shirt. He has his arm around the woman's shoulder. The woman, on the right, has short brown hair and is wearing a black t-shirt. She has her hands clasped in front of her. Both are smiling. In the foreground, there are two glasses of beer on the table. The background is a dark brick wall.

WCRE 2010
Boston, MA, USA



Vadim Zaytsev @grammarware · Feb 5, 2014



The best “best paper award” talk by Andi Marcus ever.
Exposes the whole life cycle of rejections of the paper!

[#csmrwcre](#) [#saner](#)

GETTING A PH.D.

WHY PURSUE A PH.D.?

Career choice

Academia – research and/or teaching

Industry – R&D

Personality

Problem solver

Like independent work

Dislike routine work

WHY PURSUE A PH.D.?

Career choice

Industry – programmer/developer

Personal

Prestige – I want to be called **Dr. X**

Everyone in my family has a graduate degree

Pathway to legal immigration

GOALS OF THE PH.D. PROGRAM

Training (acquiring skills)

Research

Organization

Teaching and mentoring

How to do research

About your research

In your field

Research work

Advance knowledge in the field

Can be done in several ways

Expertise

At graduation you will be the worlds' foremost expert in your research topic

WHAT DO YOU NEED BEFORE STARTING YOUR PH.D.?

Motivation and patience

It is an investment in your future

Communication skills

Orthogonal to language

If YOU cannot explain what you are doing, then nobody can

Stamina

Not a 40 hours/week job

Research does not stop at 5pm

Ability to learn

New things

From failures and successes

WHAT DO YOU NEED BEFORE STARTING YOUR PH.D.?

Background

Technical skills

Broader background, beyond CS/SE, is a plus

Language

English

Personality

Curiosity

Ability to work with others

Humility

Self confidence

CHALLENGES DURING THE PH.D.

Research topic

Advisor

Pressure and coping

Constant judging - papers, grants, presentations, classroom

Repeated failures

Deadlines

Balance work-life

Finances – most students are poor

WHAT IS YOUR PH.D. ALL ABOUT?

Independent research

Expertise

you will be the worlds' foremost expert in your research topic

Until that happens, remember that

“Your advisor is always right!”

Jonathan Maletic

“When you manage to prove your advisor wrong, you are ready to graduate”

Andrian Marcus

WHAT IS YOUR PH.D. ALL ABOUT?

Benefits

Exciting, sense of purpose, work with others, unparalleled satisfaction, it is work for you, it is the beginning of other things

Cost

Hard work, frustration, high failure rate

Warning!

May lead to nowhere, not for everyone, you often need to give up independence to achieve independence, it can go really wrong

Not the end of the world if you never finish your Ph.D.

WHAT IS YOUR DISSERTATION ABOUT?

The end results of years of work, sweat, and struggles

Some say you should even publish it as a book

Byproduct of your research training and work, on the way to build a career and long term research agenda

Many places accept dissertations as a collection of (some) of your papers

CHOOSING YOUR RESEARCH TOPIC

Listen to your advisor

You need their help and expertise, especially as you start your research

Do not go too far away from their area of expertise, yet
do not stay too close either

The most valuable help you can get is not technical in nature

Publication venues, research techniques, past experience, connections, etc.

CHOOSING YOUR RESEARCH TOPIC

Be brave

“Boldly go where no one has gone before”

Be prepared

It may lead to nowhere, so you need to come back to square one

Beware of “rabbit holes”

You need to sell it and protect it

Be mindful about return on investment

**WHAT KIND OF RESEARCH
DO YOU WANT TO DO?**

PASTEUR'S QUADRANT

*Basic Science
and Technological
Innovation*

Donald E. Stokes



Quest for fundamental understanding

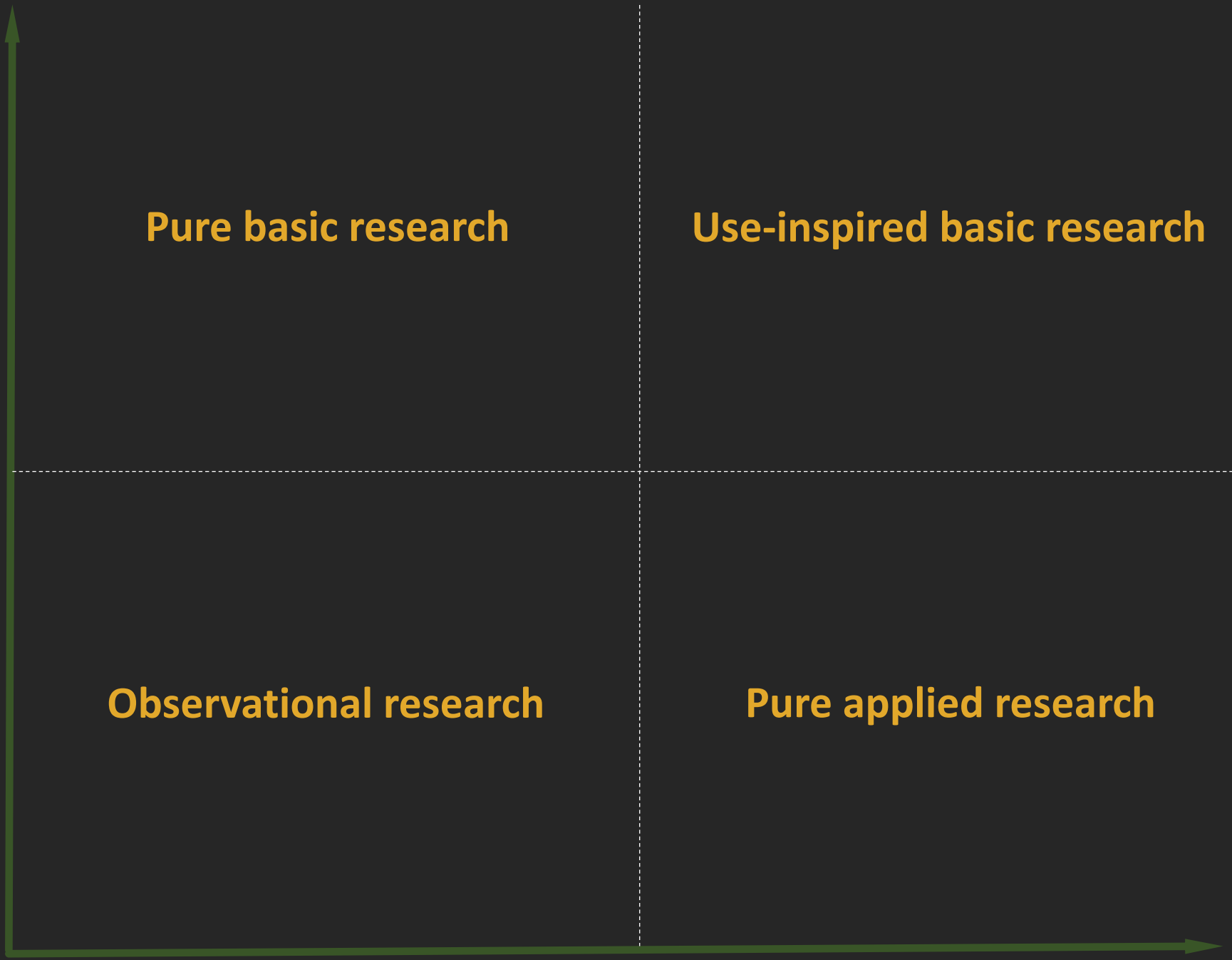
Pure basic research

Use-inspired basic research

Observational research

Pure applied research

Consideration of use



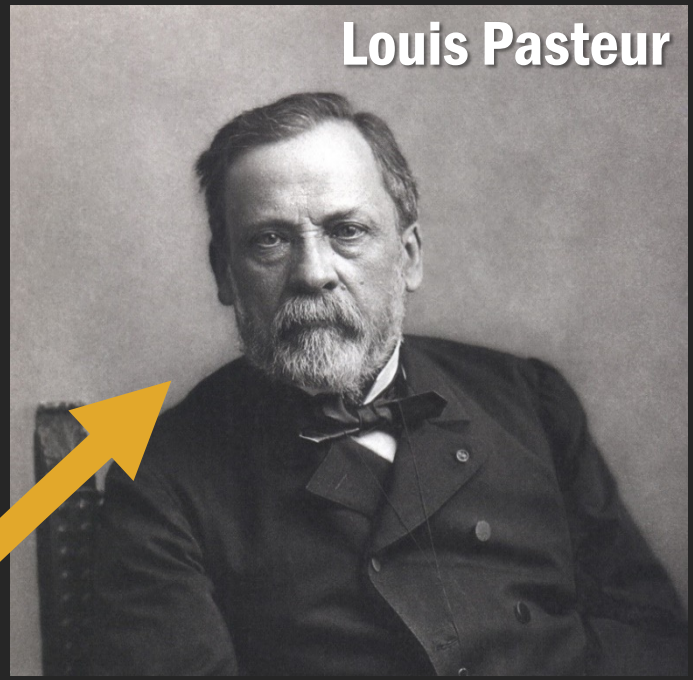
Quest for fundamental understanding

Pure basic research

Niels Bohr



Louis Pasteur

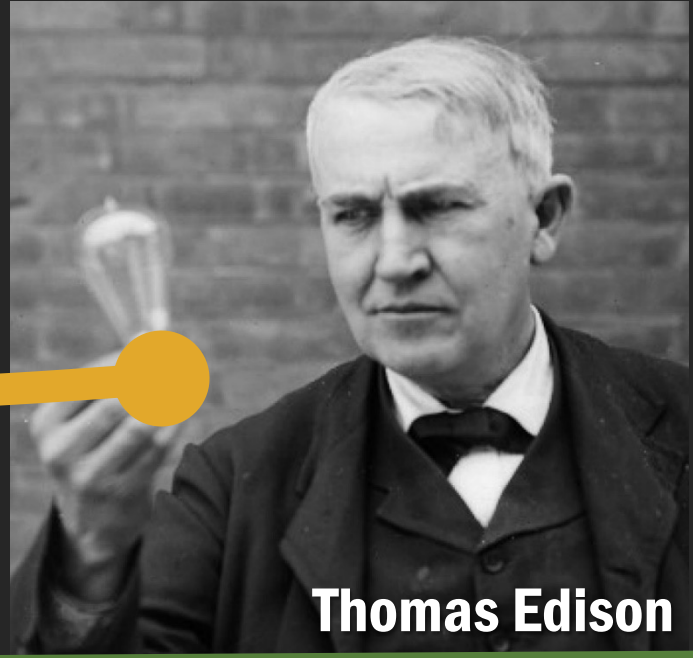


Use-inspired basic research

Observational research



Birdwatchers



Thomas Edison

Pure applied research

Consideration of use



RESEARCH IN SOFTWARE ENGINEERING

UNDERSTANDING HOW TO AND HELPING

produce software cheaper

- increase productivity
- process improvement
- tool support:
IDEs, programming languages,
algorithms, libraries, etc.
- usability
- communication
- information management
- cognition
- human factors

produce better software

- software quality analysis and
measurement
- software testing
- process improvement
- tool support:
IDEs, programming languages,
algorithms, libraries, etc.
- usability
- communication
- information management
- cognition
- human factors

RESEARCH AREAS IN SOFTWARE ENGINEERING

Process improvement is difficult and long-term

Done in academia during the 1980s

Academic environment inappropriate for such research

Mostly grassroots movement since 1990s (e.g., agile community)

Productivity improvement research is tricky

Needs industrial environment

Requires research methods from other disciplines, such as, anthropology, social science, management, HCI, etc.

Driven by trial and error, e.g., unusual work environments

Hated by developers, loved by managers

Needs high-quality process data

Serious ethical issues

DEFINING YOUR RESEARCH AGENDA

What?

Why?

How?

WHAT?

What is the problem you are researching?

The most common question you will have to answer in your career

WHY?

Why is this important?

The most important question you will have to answer in your career

HOW? – CONDUCTING YOUR RESEARCH

The good news

There is no one way (or best way) to do research

The bad news

There are bad ways to do research

**HOW DO WE LEARN
TO DO RESEARCH?**

LEARN FROM THE EXPERTS



Embarking On a Journey to Conduct Disruptive Research in Software Engineering: *Who, What, How*

Together, we will take a deep dive and follow a less traveled path through the landscape of research methods in software engineering research. We will explore *who* our research aims to impact, *what* kinds of contributions we can expect from our research, and *how* we can use innovative research methods. Some of the topics we will dive into include **design science** as a frame for software engineering research, the benefits and challenges of using **mixed methods** in software engineering, and how to uncover the potential but not always obvious or positive **disruptive impacts of novel technologies** (such as generative AI and VR) on software engineering practice. After this talk, you should feel more empowered to pursue ambitious and impactful research using innovative research methods.



Jonathan Maletic
@KSU



ICPC '08
Amsterdam, NL

Václav Rajlich
@WSU

ICSM'05
Philadelphia, PA



Max Di Penta
@Uni Sannio

ASE'17
Singapore



"If all you have is a hammer, everything looks like a nail." – Abraham Maslow



STUDY AND UNDERSTAND THE PROBLEM



IDENTIFY SPECIFIC INSTANCES YOU CAN SOLVE



FIND THE BEST SOLUTION



RESEARCH IN THE SEERS group AND CO.

Need driven research

(as opposed to hypothesis driven)

Empirical research

(as opposed to theoretical)

1. Identify an existing problem
2. Document the problem
3. Study and understand the problem
Add to the knowledge in the field
Transfer knowledge to other problems – future research
4. Formulate a solution
5. Implement the solution
6. Evaluate the solution empirically
7. Publish the results
8. Can we generalize the solution?
9. Can we adapt the solution to other similar problems?



BREADTH VS. DEPTH

How far should you explore a problem?

Novelty wears off quickly

Being the first or one of the many

Learning takes time – expertise has costs

Adoption and impact take years (sometimes decades)

Focusing on one problem will leave room for others to take on the others

REINVENTING THE WHEEL

Very common

It starts with new terminology

Often needed

Often an excuse to ignore previous work

Buzzwords

Start with a survey

Place your work in the proper context and make the differences clear

COLLABORATIONS VS. SOLO WORK

Nobody's perfect

Others can help us

We can help others

Do a lot of small jobs with many people

Do fewer bigger jobs with few people

Issues

Intellectual property, funding, students, jobs, etc.

Cost

Working with senior collaborators

PRESENTING YOUR RESEARCH

Publishing

What?

When?

Where?

Branding

LEAST PUBLISHABLE UNIT (LPU)

Avoid LPUs!

Focus on quality over quantity

Do not ignore the world we live in (e.g., “the number game” in academia)

Graduation and tenure pressures lead to LPUs

Incremental research vs. LPU

NOW OR LATER?

Workshops are today what conferences used to be two decades ago

Fast dissemination of research

Rapid feedback from the community

Mark your territory

arXiv

HOME SWEET HOME

Find your home venue, where *everybody knows your name*

Higher chance of relevant reviews

Reviewers more likely to know your work (helps avoiding LPUs)

Cultivate your presence there and maintain it for the long term

Invest in your home – make it better

Aim for the high impact venues (ICSE, ASE, FSE, TSE, TOSEM), but beware that they are everyone's home

BUILDING A BRAND NAME

Are you an expert in the problem or in the solution?

Example: using LLMs in software engineering

Solutions are applicable across fields

Problems are within (sub)fields

Who (persons or venues) are sensible to the problem and can appreciate your solution?

ETHICS AND RESPONSIBILITIES

Everybody talks and knows about it

We do not know or talk enough about it

Largely self controlled!

Poor community infrastructure for control

YOU, YOUR ADVISOR, AND YOUR RESEARCH

Your Ph.D. research should be long term

Where does their work ends and yours starts?

Split/share the world after your graduation

Will you be competitors or collaborators after your graduation?

Who has the IP of your dissertation work?

CAREERS IN ACADEMIA

WHY SHOULD YOU WORK IN ACADEMIA?

Not an easy sell

Alternatives are often more glamorous or lucrative

You probably heard your advisor complain about academia every day for the last x years

WHAT I LIKE ABOUT IT?

More than a job

“Choose a job you love, and you will never have to work a day in your life.”

- unknown - popularized by Harvey Mackay

Five jobs into one

Educator; researcher; mentor; administrator; promoter

Independence

Academic freedom - tenure

Unique perks

Work with young and smart people

World wide collaborations + travel

Flexibility time-wise and work-wise

Sabbatical

WHAT I LIKE ABOUT IT?

Stability

Tenure + Good salaries

Consulting + start-ups

Research impact

Publications - citations

Consulting - industry

People impact

“If you think in terms of a year, plant a seed; if in terms of ten years, plant trees; if in terms of 100 years, **teach the people.**” – *Confucius*

Legacy

Students – careers

Mentees – school of thought

MYTHS ABOUT WORKING IN ACADEMIA

I have to teach a lot

Most research active faculty will teach 2 courses a year

I will not make money

True in some countries

Can have research income, consulting, business

Research is too hard

There are teaching based positions

Can migrate towards administrative positions

Research is not practical

It can be

LEGACY

ICSME'11
Williamsburg, VA



LEGACY



ICSME'16
Raleigh, NC

LEGACY



ICSME'16
Raleigh, NC