



# LADIES AND GENTLEMEN, START YOUR ENGINE!

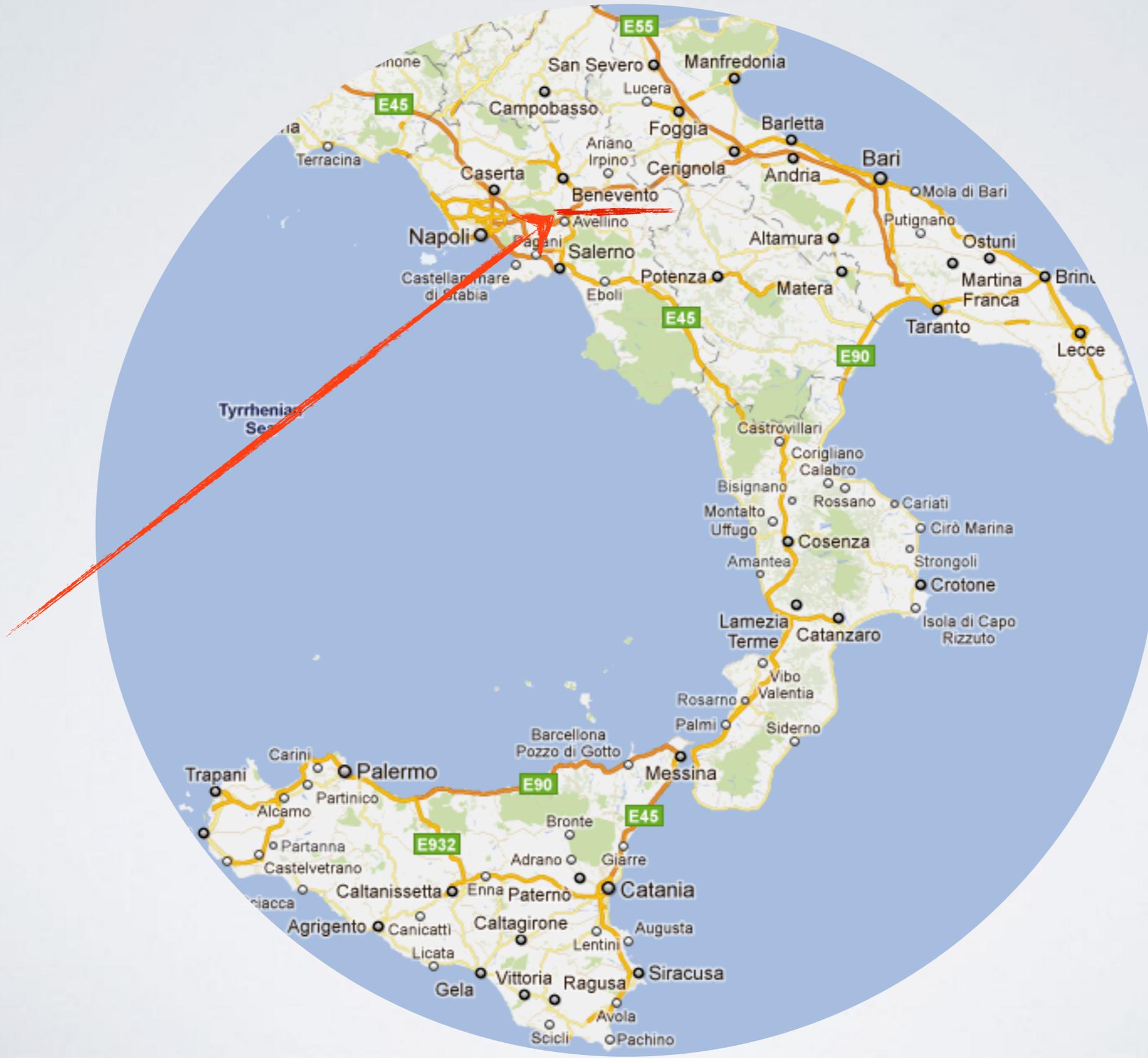
Being Successful with your PhD  
Without Crashing at the First Corner

**Massimiliano Di Penta**

University of Sannio, Italy

[dipenta@unisannio.it](mailto:dipenta@unisannio.it), [@maxdipenta](https://www.instagram.com/maxdipenta)

# UNIVERSITY OF...WHAT?



# ABOUT ME

Teaching Advanced Software Engineering (Master)  
+ an introductory course (Bachelor)  
+ empirical software engineering (PhD)

Chair of the computer science engineering  
Bachelor and Master program

I review A LOT

# MAIN RESEARCH INTERESTS

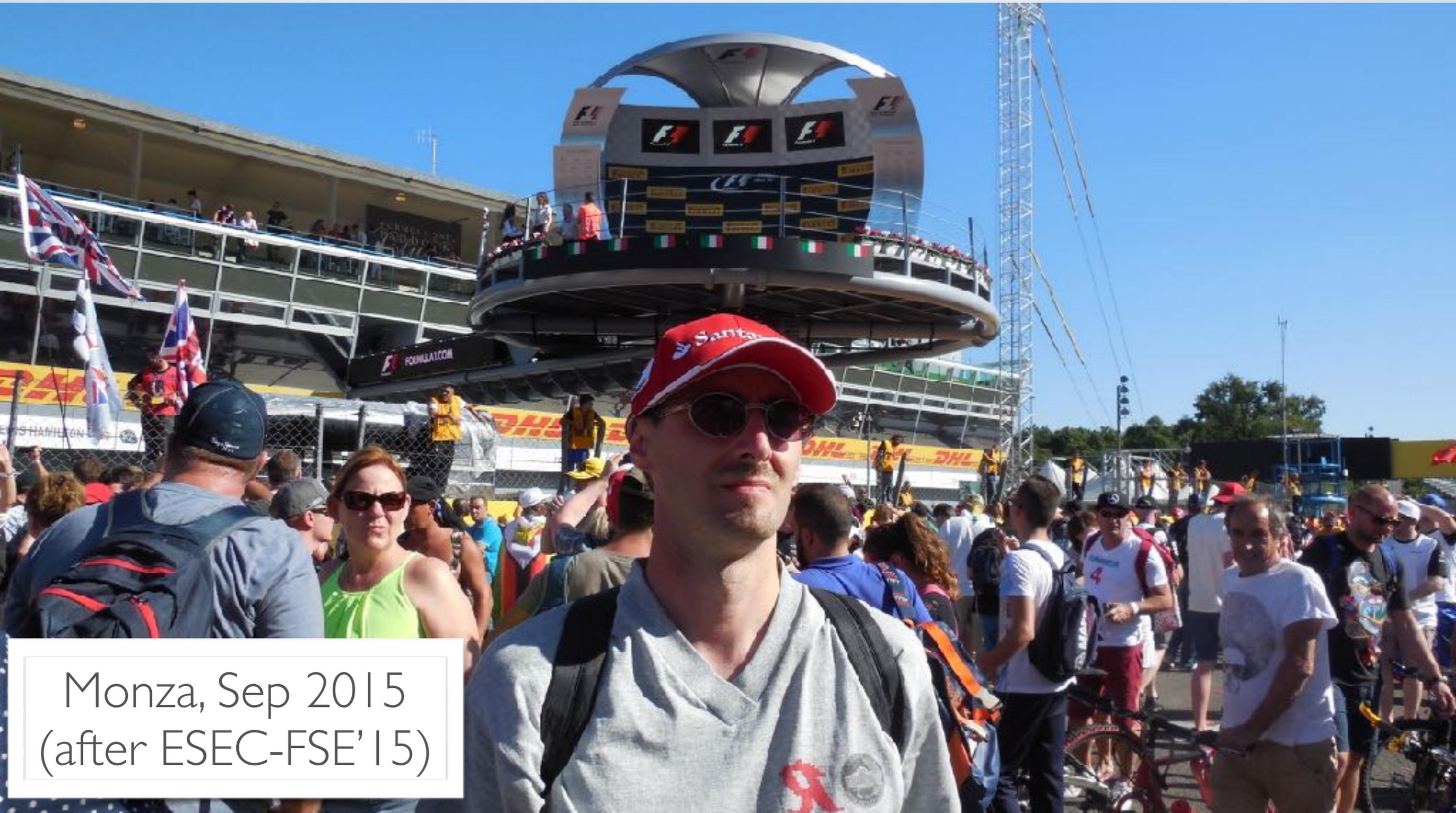
Software evolution

Software analytics

Empirical software engineering

Software testing

# OTHER INTERESTS



Monza, Sep 2015  
(after ESEC-FSE'15)

# OUTLINE



Conducting  
your Research

Managing  
your PhD

Contingency  
Plans

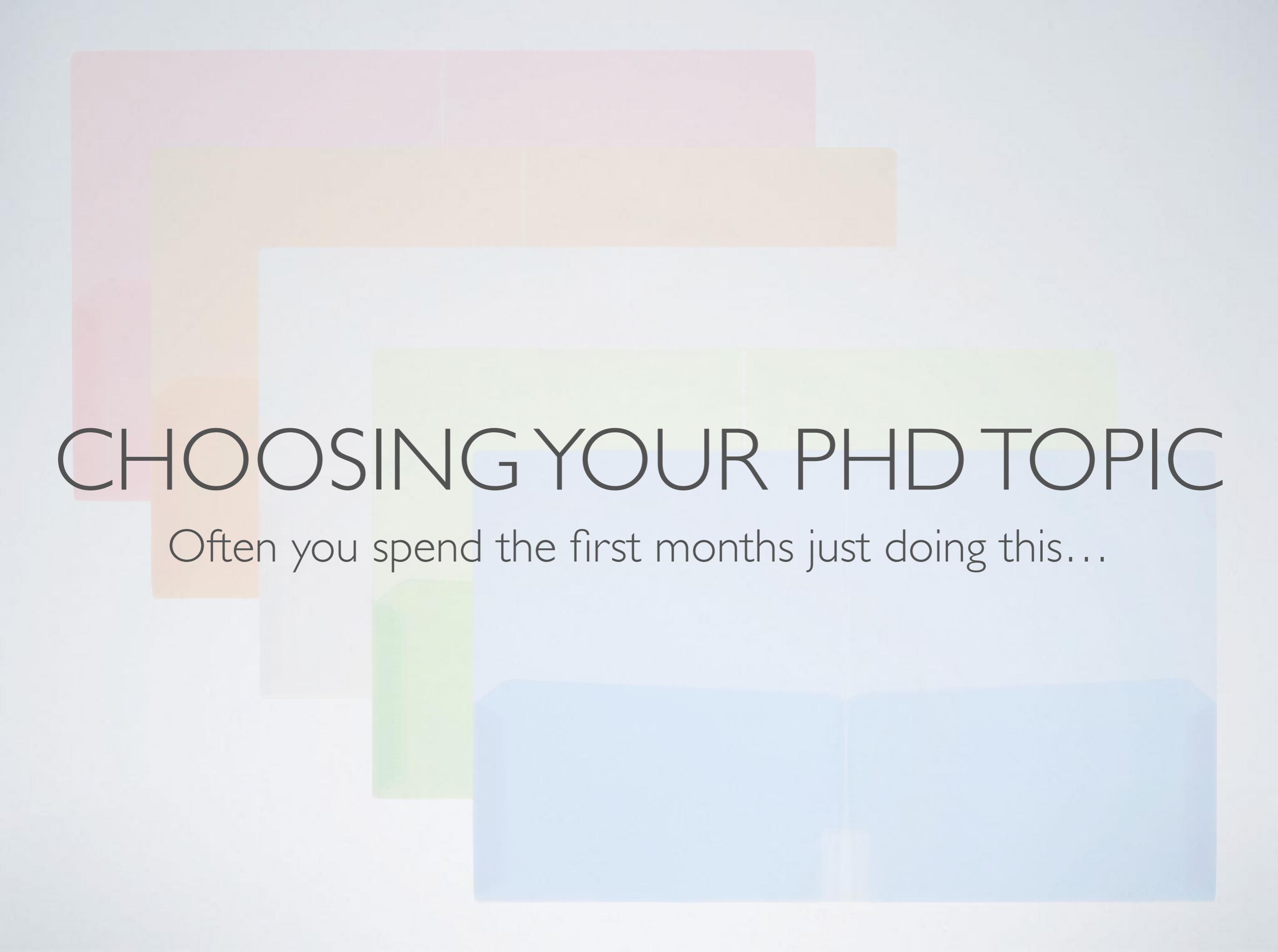
Next Steps

LET'S START!



# CONDUCTING YOUR RESEARCH





# CHOOSING YOUR PHD TOPIC

Often you spend the first months just doing this...

# THE TOPIC

- (Still) relevant in research and practice
- You feel there is room to contribute
- Your research group has competences
- **You like it**

# GOOD RESEARCH

Solves a real problem practitioners have

Builds something useful for future researchers

Makes observations from which others can learn

# PITFALLS (IMHO)



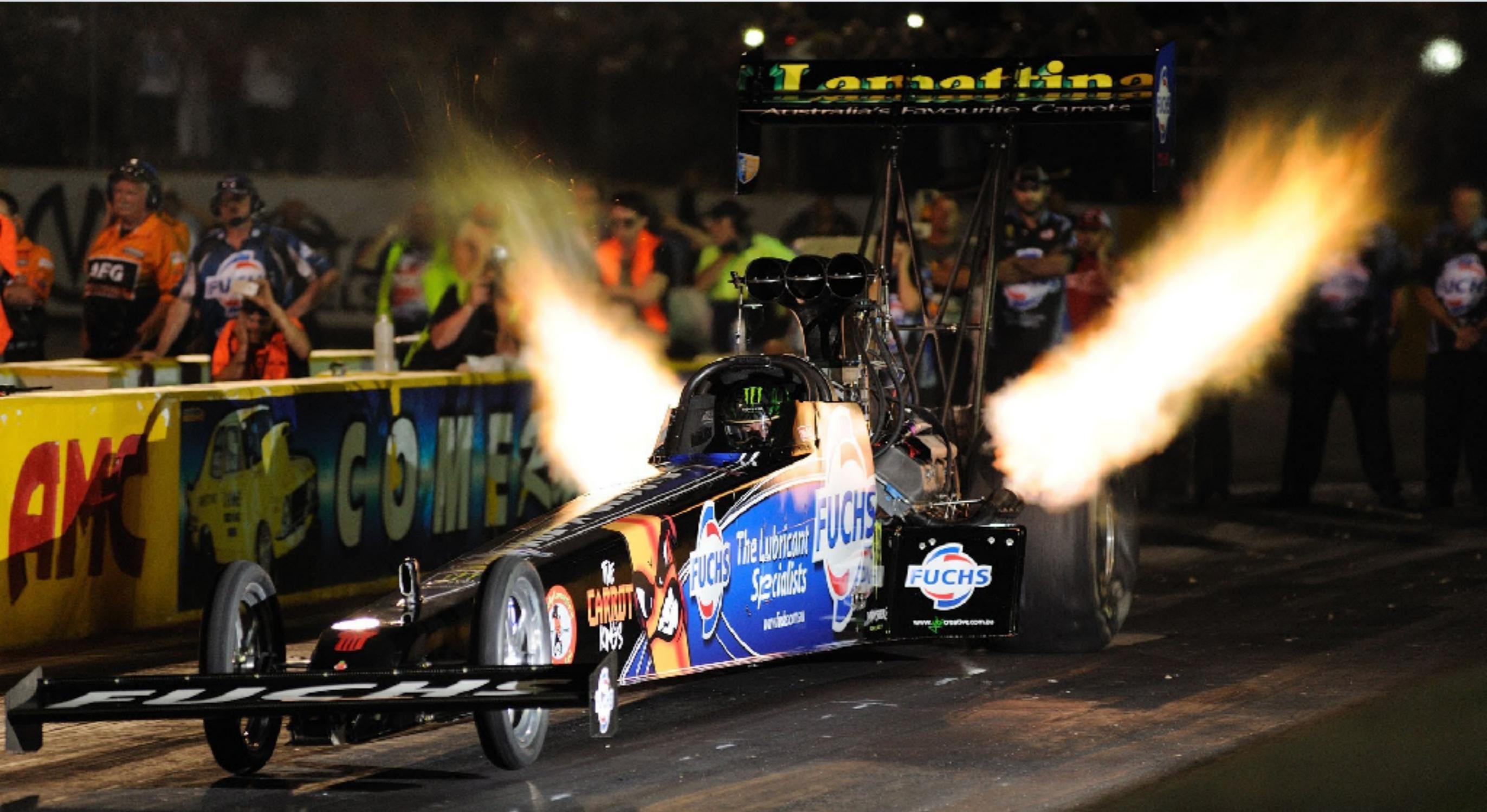
WE HAVE A SOLUTION, LET'S  
FIND A PROBLEM FOR IT



# SALAMI PAPERS



# RUSHING TO THE NEXT DEADLINE



HOWEVER....

# RACE STARTS AT 2PM



Conference submissions close at Midnight AOE  
(2pm CET)

# PLANNING IS ESSENTIAL

## SHANGHAI - PIT STOP SUMMARY

AVAILABLE  
TYRES



WHITE MEDIUM



YELLOW SOFT



RED SUPERSOFT



GREEN INTERMEDIATE



BLUE WET



# PARK FERME' (CLOSED PARK)



Refrain to perform massive changes

- New computations
- Major manuscript reorganization

# BE PICKY

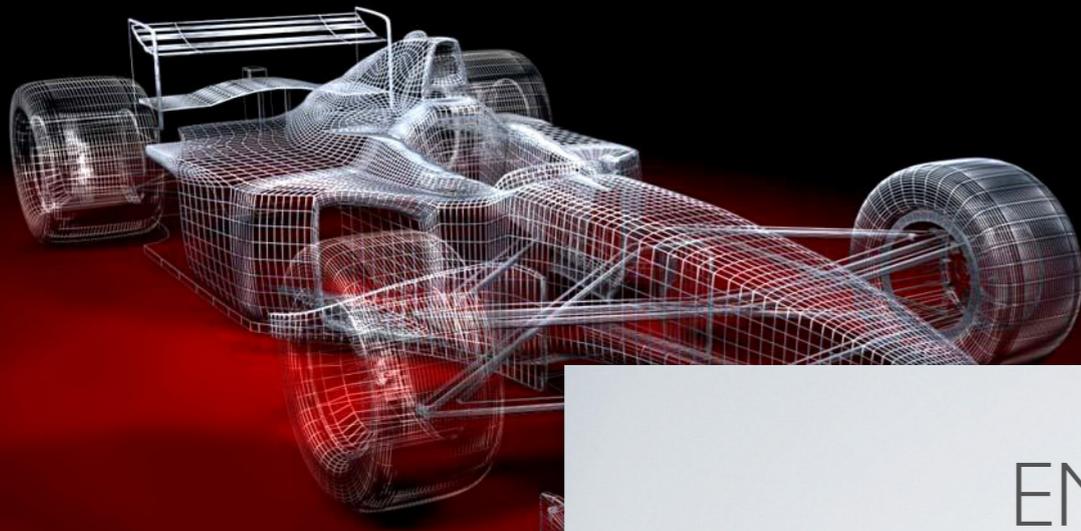
Proof read, proof read, proof read

Avoid typos

Polish any aspect of the paper (bib, figures, tables)

# DIFFERENT CONTRIBUTIONS

THEORETICAL



PRACTICAL

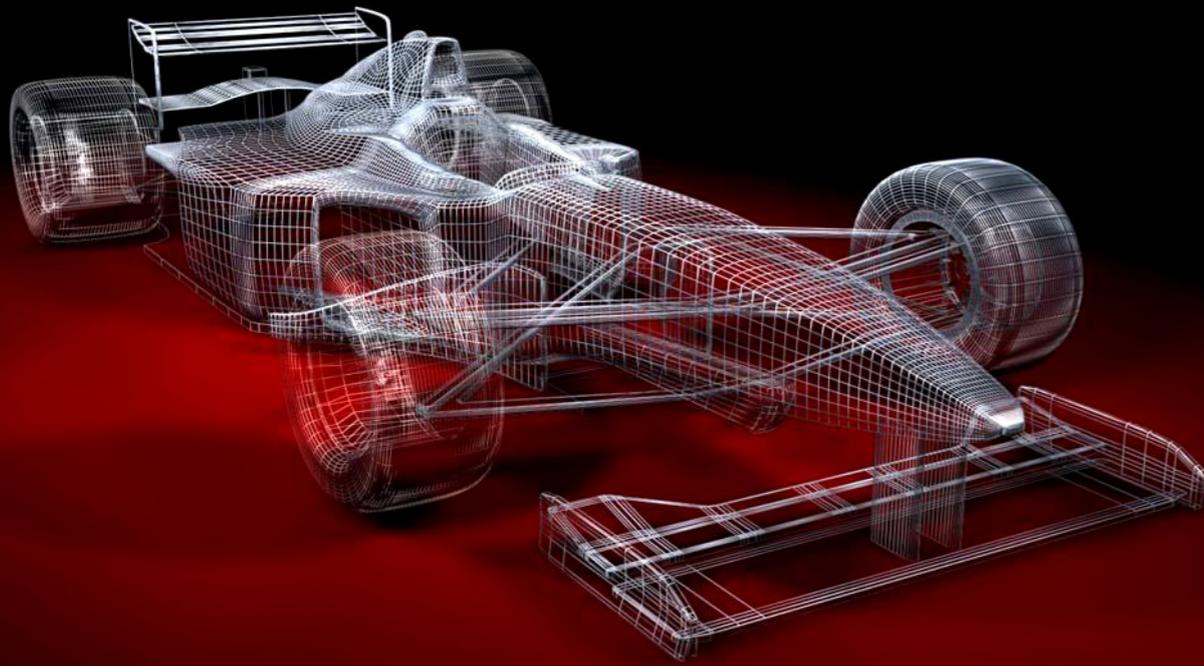


EMPIRICAL



# DIFFERENT CONTRIBUTIONS

THEORETICAL



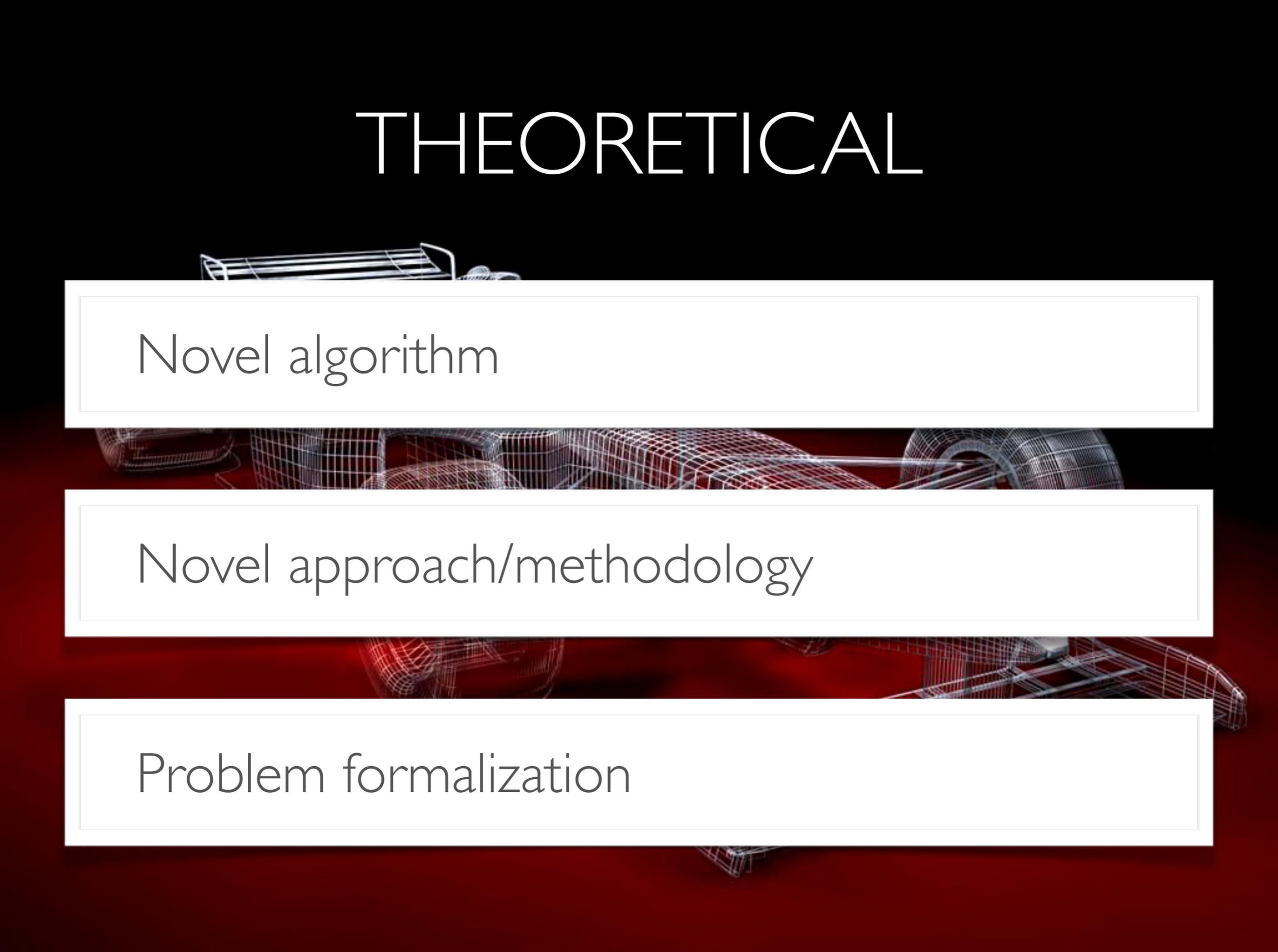
PRACTICAL



ICAL



# THEORETICAL

The background of the slide features a dark red surface with several white wireframe models of various 3D objects, including spheres, cylinders, and rectangular prisms, scattered across it.

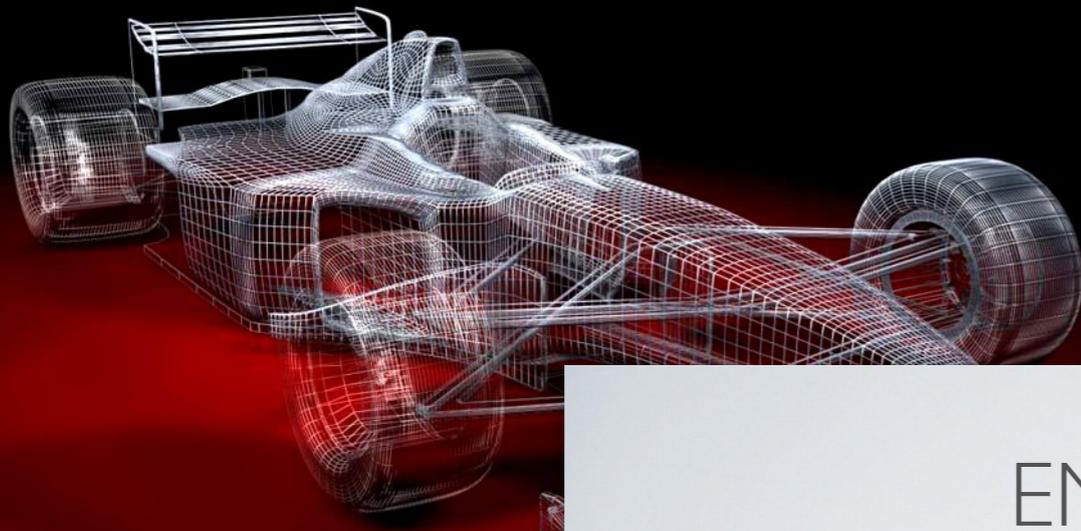
Novel algorithm

Novel approach/methodology

Problem formalization

# DIFFERENT CONTRIBUTIONS

THEORETICAL



PRACTICAL

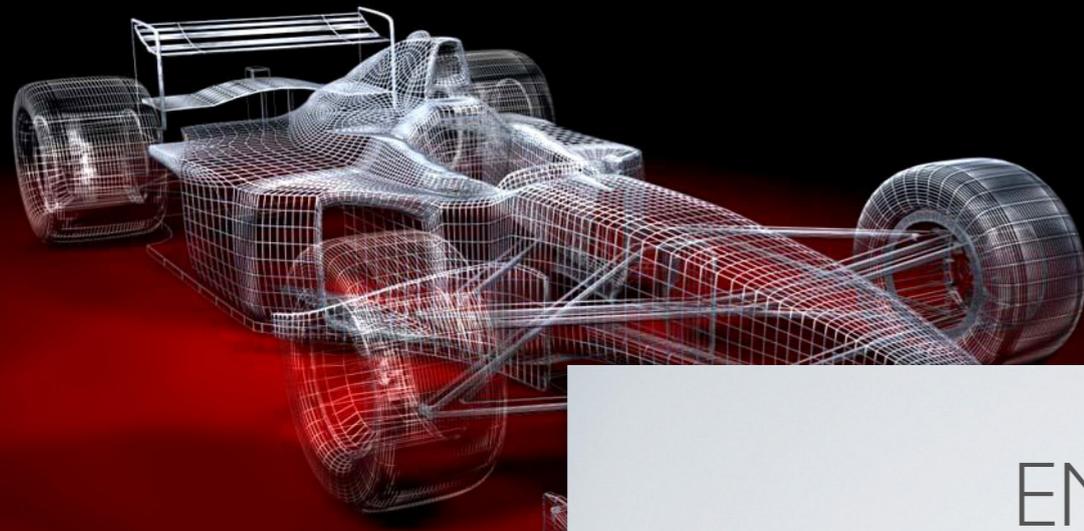


EMPIRICAL



# DIFFERENT CONTRIBUTIONS

THEORETICAL



PRACTICAL



EMPIRICAL



# PRACTICAL

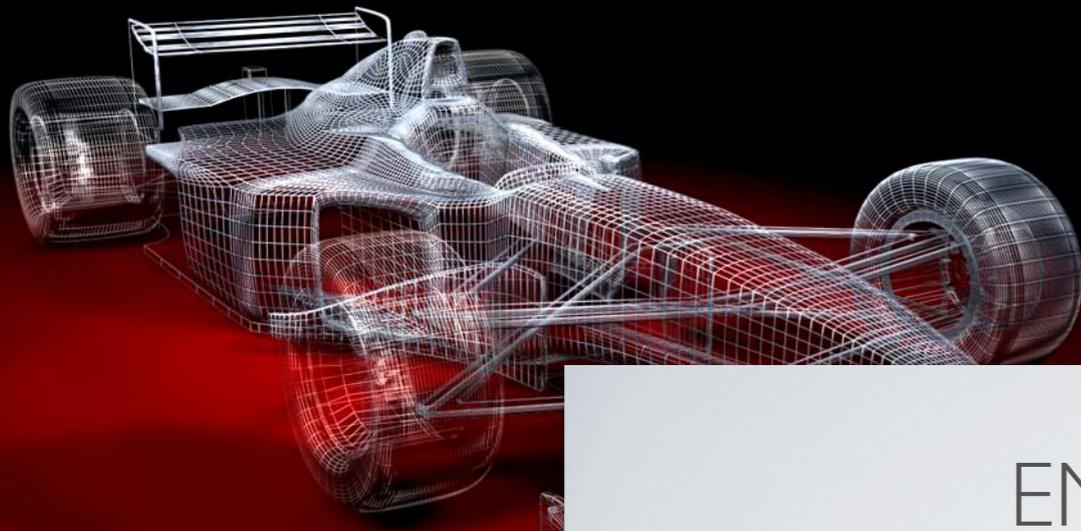
The background image shows a workshop environment where several people are working on a car chassis. The car is white with orange accents and is mounted on a metal frame. The people are wearing red jackets, and one person's jacket has 'KWCCT' written on it. The scene is brightly lit, and the focus is on the car and the workers.

(Usable) tool

Dataset

# DIFFERENT CONTRIBUTIONS

THEORETICAL



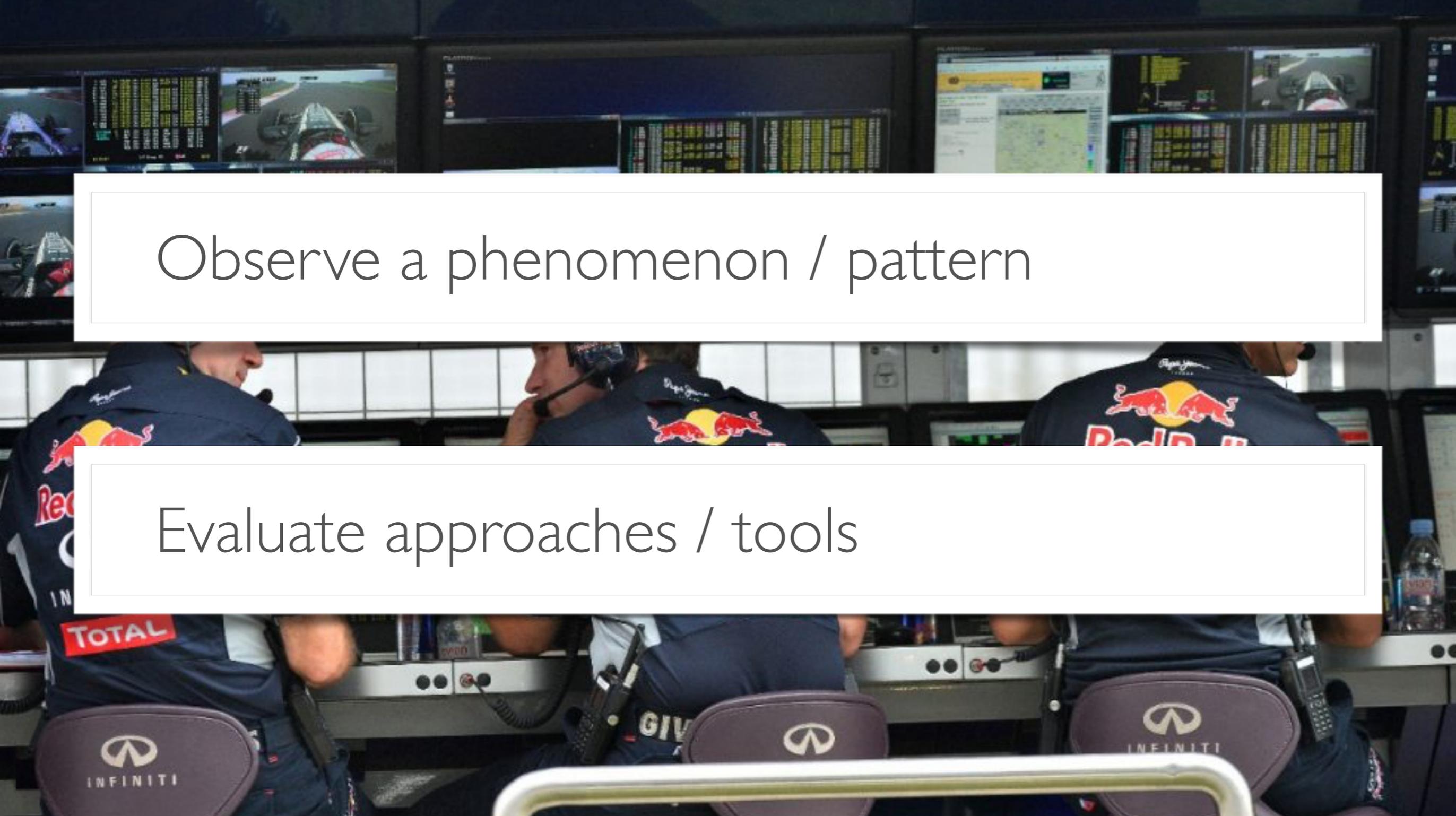
PRACTICAL



EMPIRICAL



# EMPIRICAL

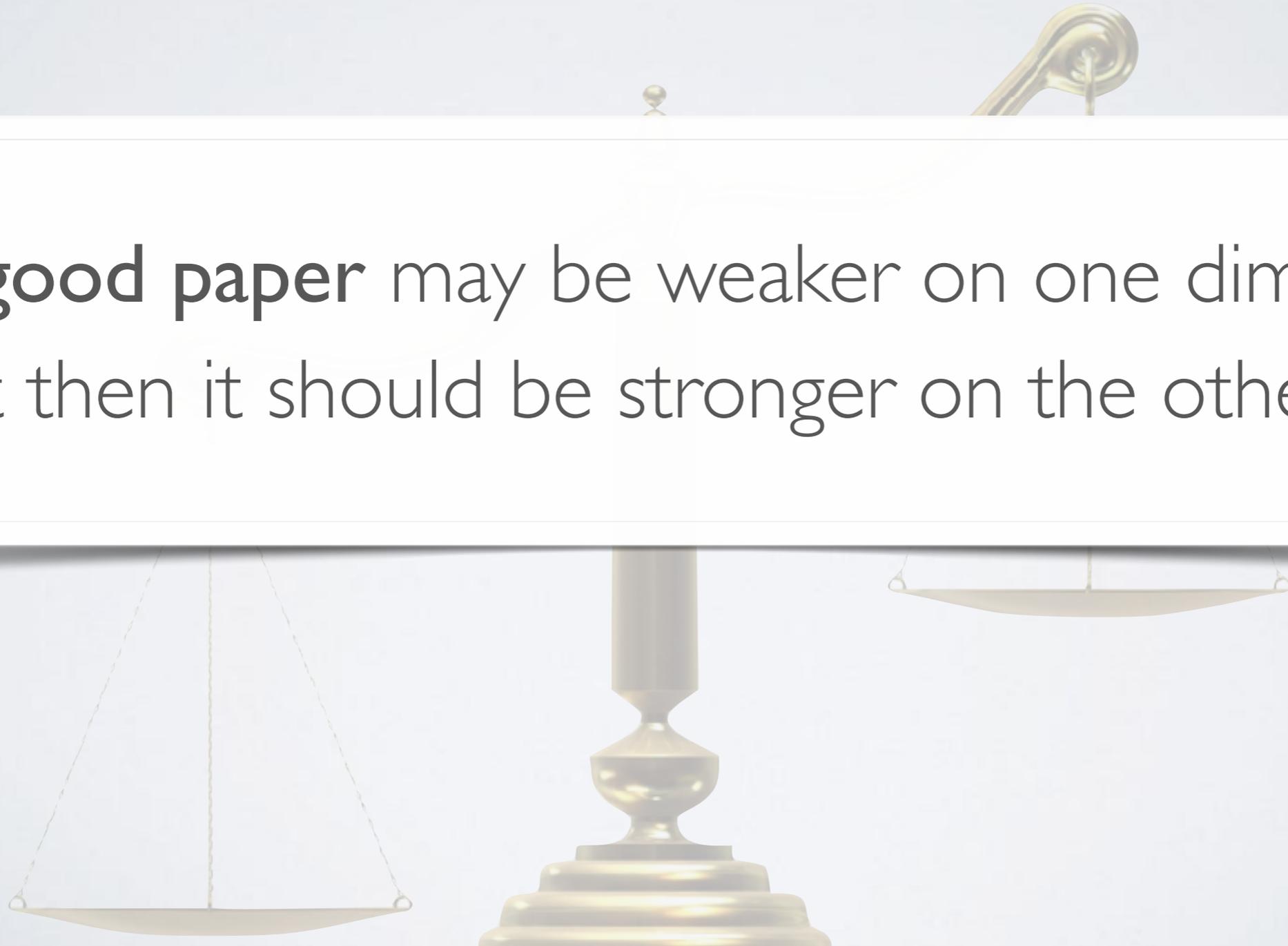


Observe a phenomenon / pattern

Evaluate approaches / tools

# HINTS

## (FOR AUTHORS AND REVIEWERS)



A **good paper** may be weaker on one dimension, but then it should be stronger on the other side

# WHAT DOES IT MEAN?

- Your approach provides a small theoretical increment
- You should provide a strong empirical evidence that your approach is better than the current state of the art

# CHALLENGES

- No tool available for alternative approaches
- No data
- The paper does not contain enough detail to reimplement the approach
- Asked the authors, but they don't respond

# MONACO '84: A RELATIVELY MINOR CAR AND A GREAT DRIVER...



# THE RESEARCH PAPERS I LIKE

A **great paper** contributes along all dimensions

# THE IDEAL PATH

1. You conceive a novel approach
2. You evaluate the approach
3. You actually implement it in a usable / downloadable tool
4. You show that the tool actually helps



Even if you don't have all these in a paper, it would be great to have them in your thesis

# EMPIRICAL



# EVALUATING YOUR RESEARCH



# TYPICAL QUESTIONS YOU ASK

# HOW ACCURATE IS IT?



# HOW FAST IS IT?



IS IT ANY BETTER THAN  
COMPETITOR'S TOOL?



WHAT ARE WE MISSING HERE?



# RECOMMENDER SYSTEMS

“A software application that provides information items estimated to be valuable for a software engineering task in a given context”

Martin P. Robillard, Robert J. Walker, Thomas Zimmermann:  
Recommendation Systems for Software Engineering. IEEE Software 27(4): 80-86 (2010)

# RECOMMENDER SYSTEMS

“A software application that provides information items estimated **to be valuable for a software engineering task in a given context**”

Martin P. Robillard, Robert J. Walker, Thomas Zimmermann:  
Recommendation Systems for Software Engineering. IEEE Software 27(4): 80-86 (2010)

IS THE TOOL GOING  
TO HELP A DEVELOPER  
FOR A GIVEN TASK?

DOES IT WORK IN MY  
SPECIFIC COMPANY?

# DIFFERENT KINDS OF STUDIES



[Linkerman and Rombach, 2000]



# TOOL (PROMPTER)

The screenshot displays a development environment with several windows:

- Code Editor:** Shows a Java method `unZipIt` for unzipping a file. The code includes comments and logic for creating an output directory and reading the zip file content.
- Notification Center:** Contains two notifications with progress indicators. The first notification, titled "Java ZIP - how to unzip folder?", is marked with a red circle containing the number "1" and shows a 66% progress bar. The second notification, titled "How to add a progress bar?", shows a 61% progress bar.
- Stack Overflow Document:** A document titled "Java ZIP - how to unzip folder?" is shown. It includes a rating bar and a question: "Is there any sample code, how to particaly unzip folder from ZIP into my desired directory? I have read all files from folder 'FOLDER' into byte array, how do I recreate from its file structure?". The question is marked with a red circle containing the number "2".

```
/**
 * Unzip it
 * @param zipFile input zip file
 * @param outputZipFile output folder
 */
public void unZipIt(String zipFile, String outputFolder){

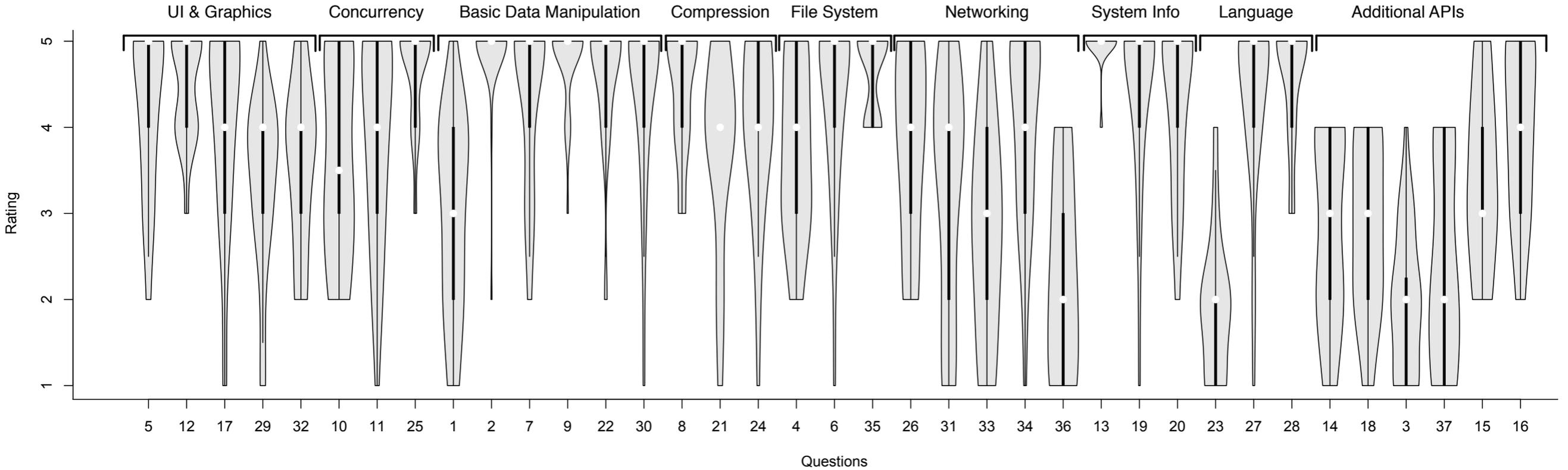
    byte[] buffer = new byte[1024];

    try{

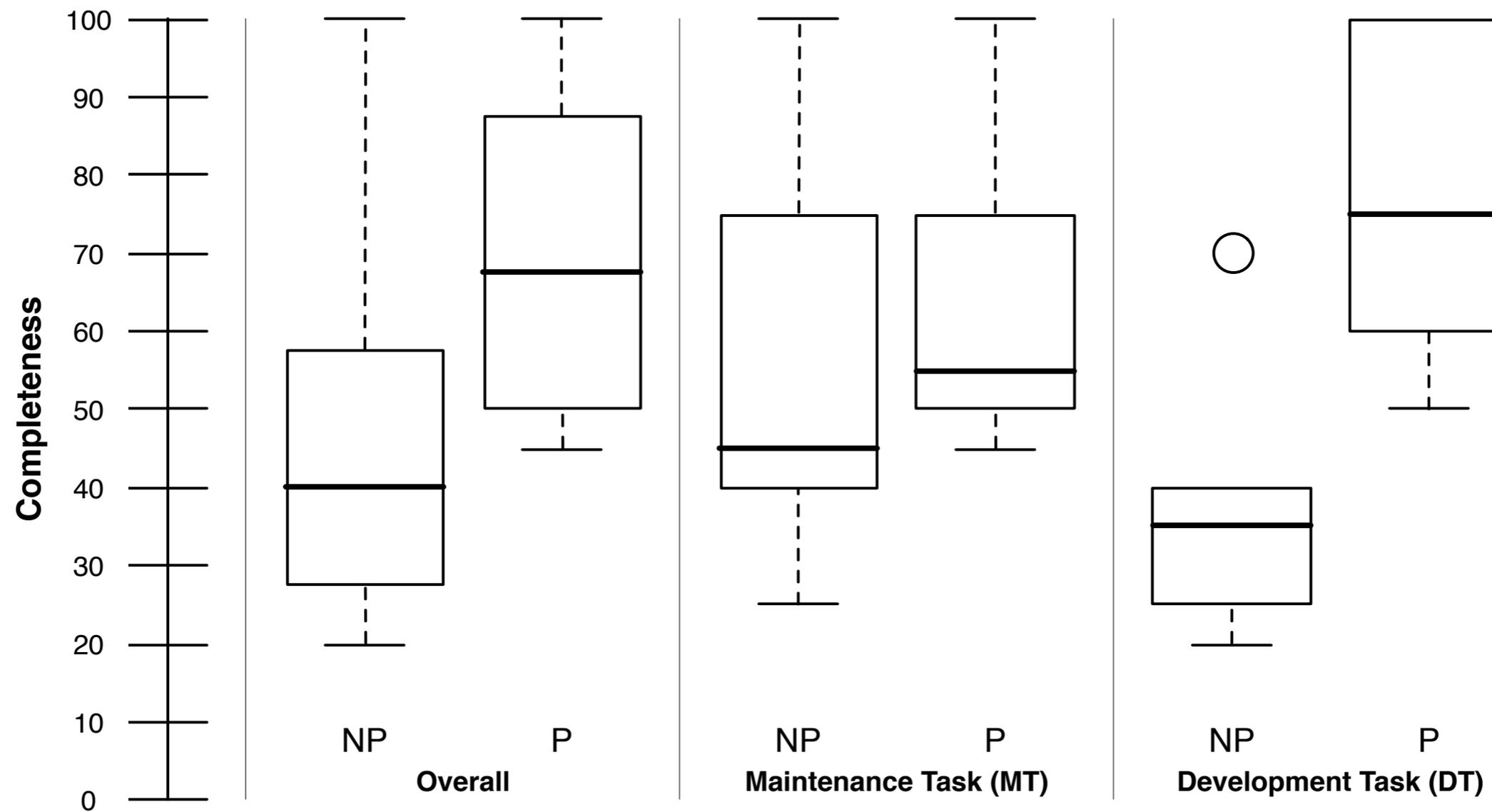
        //create output directory is not exists
        final File folder = new File(OUTPUT_FOLDER);
        if(!folder.exists()){
            folder.mkdir();
        }

        //get the zip file content
        ZipInputStream zis = new ZipInputStream(new FileInputStream(zipFile));
        //get the zipped file list entry
```

# ARE RECOMMENDATIONS RELEVANT?



# DOES IT HELP?



# MANAGING YOUR PHD



# THE ADVISOR



# DON'T EXPECT FROM YOUR ADVISOR

- Dealing with low-level details
- Carrying out your research
- **Deciding what is good for your future**

# EXPECT FROM YOUR ADVISOR

- Regular feedback
- Hints, also from a strategical point of view
- Some seminal ideas...
- Decreasing contributions along your PhD

NOT GOING TO HAPPEN



Advisor

YOU ARE THE DRIVER!



# COLLABORATIONS



# COLLABORATIONS



Computer science can be seen as a very theoretical discipline where individual work dominates...

GREAT ENGINEERING IS A  
RESULT OF TEAMWORK



# REASONS TO LIKE COLLABORATIONS

- Exchange of ideas, working habits, culture
- Opportunity to shine
- Network building
- Different competencies combined together
- Joining the force for human-intensive work

# SUGGESTIONS

**Be flexible:** different people have different working habits

**Be open** to others' ideas and opinions

# WARNING



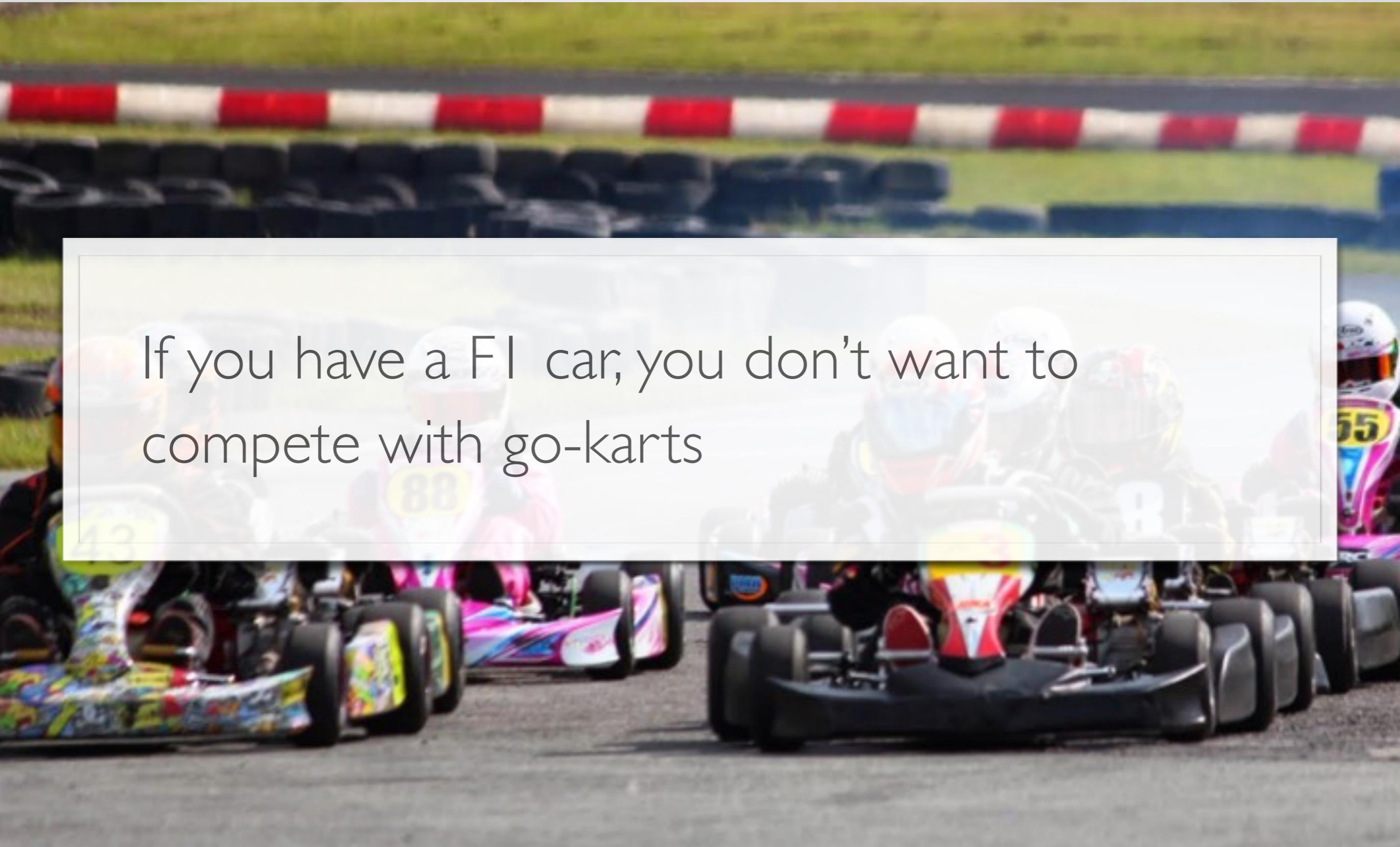
- Make sure your contribution is clear
- You need to be the leader in some of your works

# SET HIGH STANDARDS



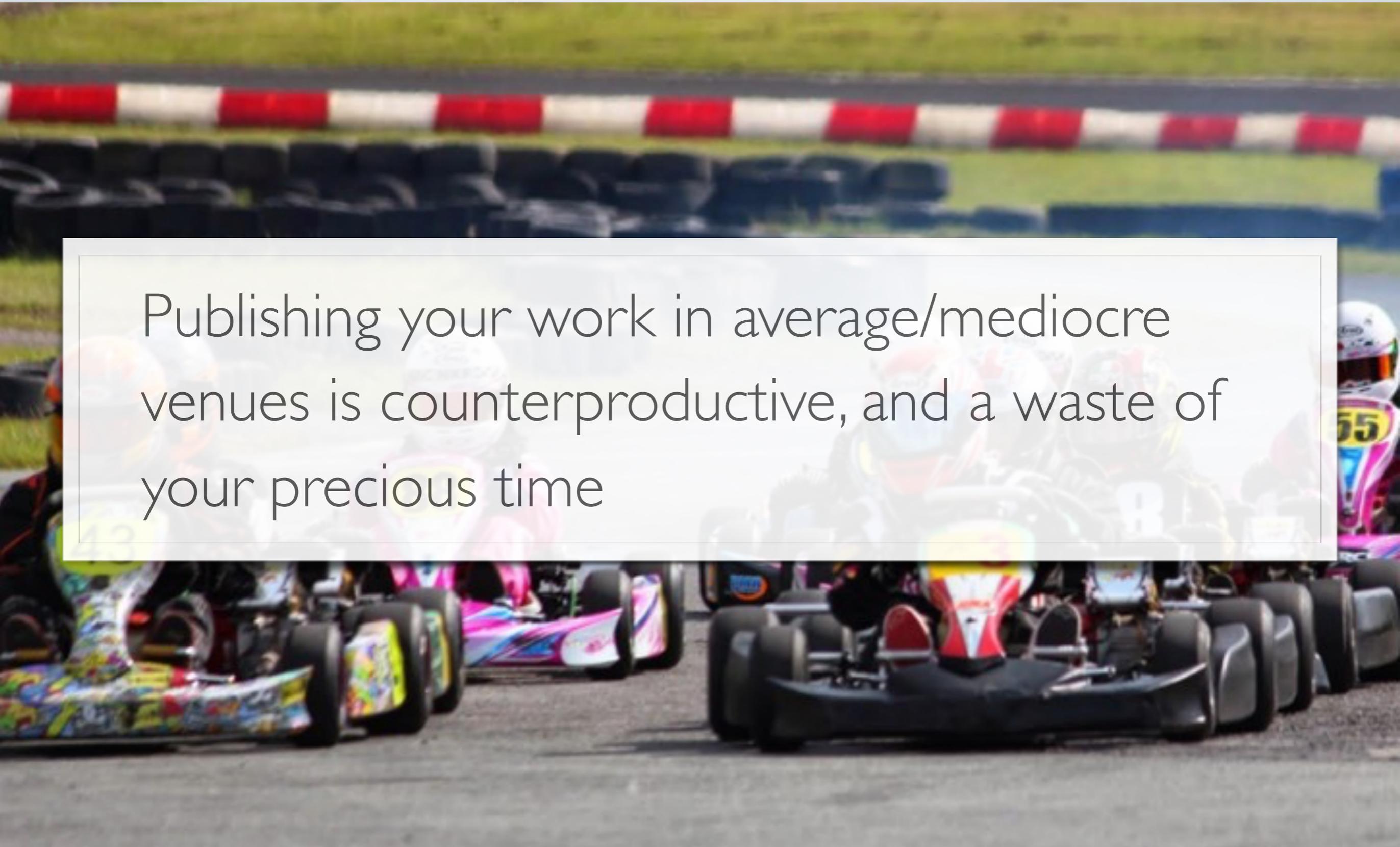
# SET HIGH STANDARDS

If you have a F1 car, you don't want to compete with go-karts



# SET HIGH STANDARDS

Publishing your work in average/mediocre venues is counterproductive, and a waste of your precious time



# WHAT ARE THE BEST VENUES?

Ask **yourself** since day zero

Always target top venues at first



# DIFFERENT OPTIONS

- **Top mainstream** journal and conferences
- **Specialized** good conferences
- **Workshops** are good for discussing ideas, better than mediocre conferences

# CONFERENCES VS JOURNALS

- Timeliness
- Visibility(\*)
- You have a deadline

- Universally recognized
- Plenty of space
- Multiple revision rounds

(\*) Now also true for journals (journal-first option)

# BE AWARE



- Different institutions/countries → different evaluation standards
- Top conference vs. high impact factor journals
- A mix of both doesn't hurt

# CONTINGENCY PLANS



# RESEARCH VS PRODUCTION



- Research is not straight-forward
- You could spend months on a solution that reveals to be wrong or inefficient



PETERSON

FIRST NATIONAL CITY  
TRAVELERS CHECKS

GOODYEAR

elf 3

elf

DON'T BLAME YOURSELF NOR  
YOUR ADVISOR!



# HINTS

When planning a piece of research, always foresee a contingency plan

Ultimately, contemplate the possibility of backtracking and rerouting

# NEGATIVE RESULTS

Negative results are worthwhile of being published

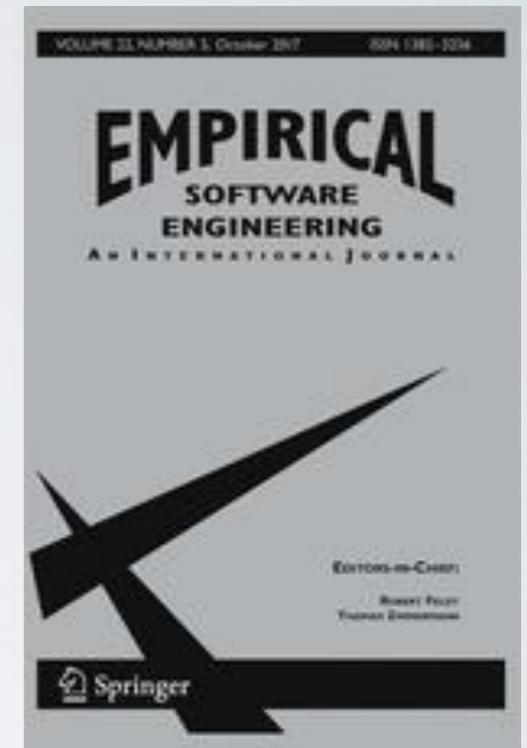
# WHAT NEGATIVE RESULTS COULD TELL US

A very successful technique completely fail in a specific context

An approach is not ready for real-world application

# SOME EXAMPLES

Special Section of the EMSE Journal,  
October 2017



<https://link.springer.com/journal/10664/22/5/page/1>

# DOES SENTIMENT ANALYSIS WORK FOR SETASKS?

## On negative results when using sentiment analysis tools for software engineering research

Robbert Jongeling<sup>1</sup> · Proshanta Sarkar<sup>2</sup> ·  
Subhajit Datta<sup>3</sup> · Alexander Serebrenik<sup>1</sup> 

Published online: 10 January 2017

© The Author(s) 2017. This article is published with open access at Springerlink.com

**Abstract** Recent years have seen an increasing attention to social aspects of software engineering, including studies of emotions and sentiments experienced and expressed by the software developers. Most of these studies reuse existing sentiment analysis tools such as SENTISTRENGTH and NLTK. However, these tools have been trained on product reviews and movie reviews and, therefore, their results might not be applicable in the software engineering domain. In this paper we study whether the sentiment analysis tools agree with the sentiment recognized by human evaluators (as reported in an earlier study) as well as with each other. Furthermore, we evaluate the impact of the choice of a sentiment analysis tool on software engineering studies by conducting a simple study of differences in issue resolution times for positive, negative and neutral texts. We repeat the study for seven datasets (issue trackers and STACK OVERFLOW questions) and different sentiment analysis tools and observe that the disagreement between the tools can lead to diverging conclusions. Finally, we perform two replications of previously published studies and observe that the results of those studies cannot be confirmed when a different sentiment analysis tool is used.

# SANER 2018 RENE TRACK



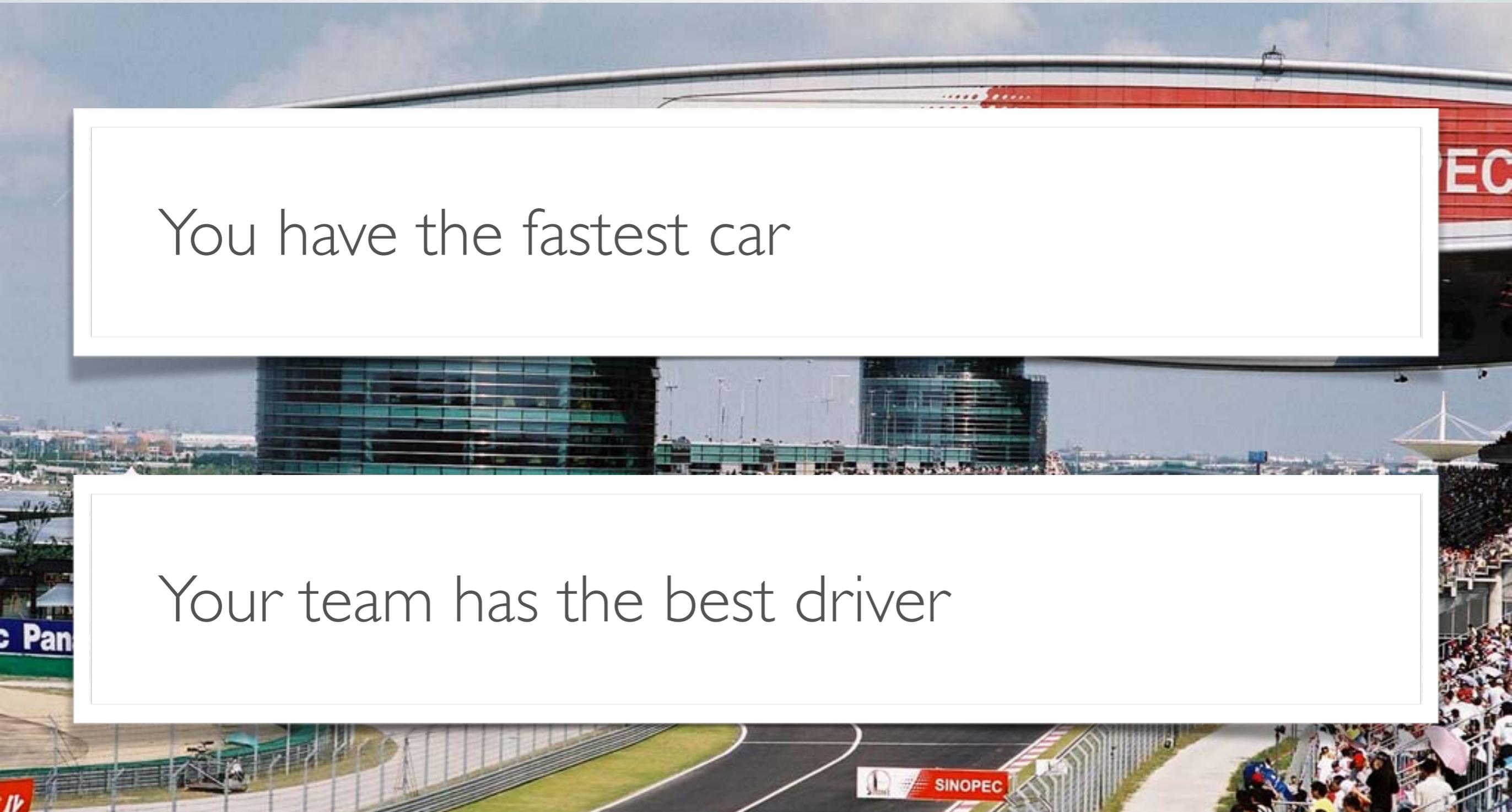
SANER 2018 <http://saner.unimol.it/>

REproducibility Studies and NEgative Results (RENE) Track

# CONFERENCE SUBMISSIONS ARE LIKE A RACE

You have the fastest car

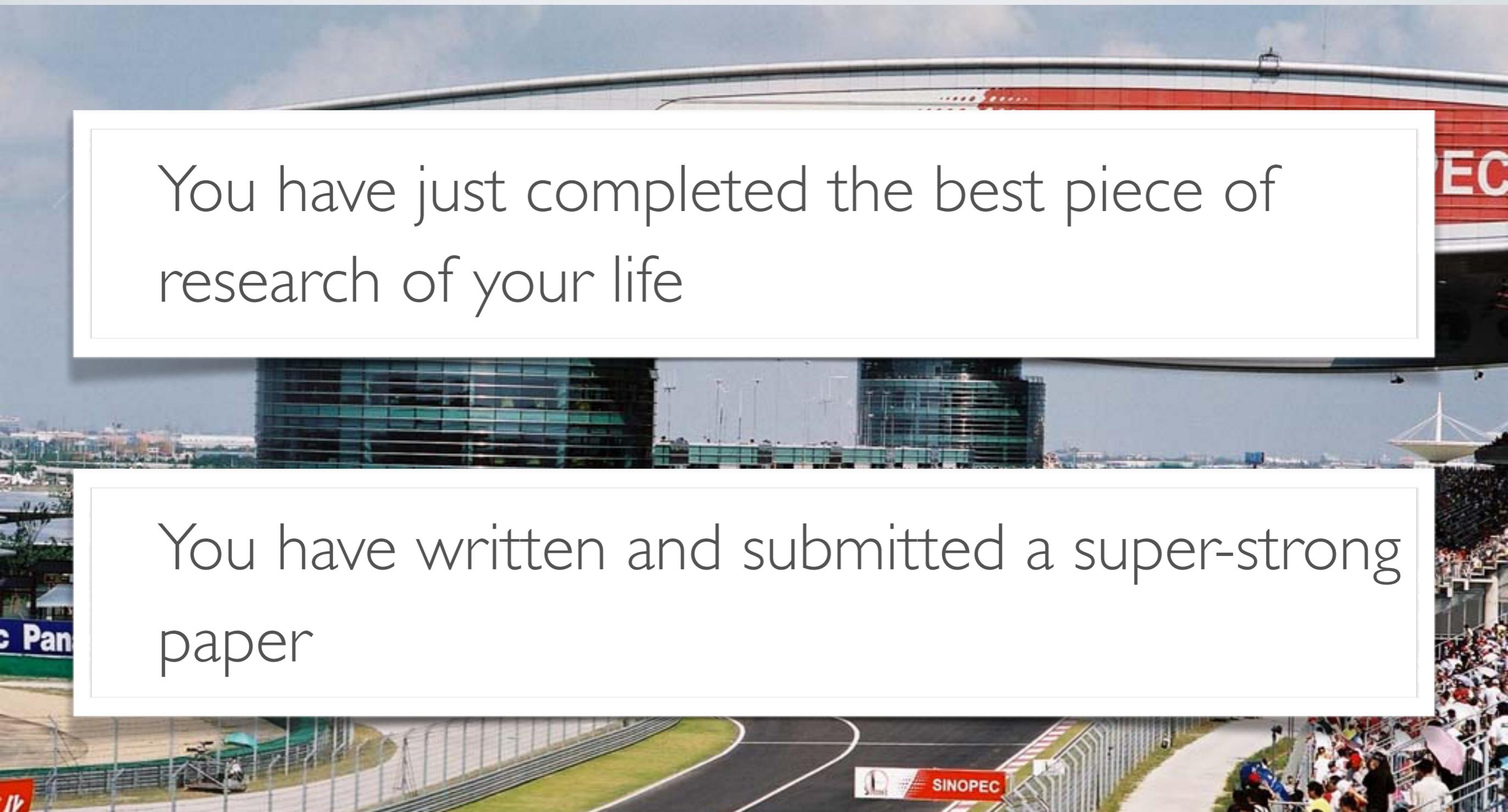
Your team has the best driver



# CONFERENCE SUBMISSIONS ARE LIKE A RACE

You have just completed the best piece of research of your life

You have written and submitted a super-strong paper



**REJECT**



# LESSONS

If your work is not good, chances to be accepted in a good venue are fairly limited

If your work is good, the unpredictable can still happen....

# WHY?

- Mild vs super-picky reviewers
- Non-expert (but still qualified) reviewers
- Having (or not) a champion or a strong detractor

# HINTS

- Never complain or even ask to reconsider
- Reviews contain main good suggestions, be critical with yourself and benefit from them
- Ok, not all suggestions might be good...

# MORE IMPORTANT

Take a rejection as an opportunity to improve your work, not as a suggestion to downgrade

YOU MAY LOOSE A RACE...



But what matters is winning the championship!

PHD ACCOMPLISHED!



WHAT'S NEXT?



# POST-DOC VS FACULTY POSITION



It really depends on what's your target and life decision

# MAIN PURPOSE OF POST-DOC



It's like Drag Reduction System (DRS) in Formula 1

# MAIN PURPOSE OF POST-DOC



- Gives you the boost to compete for a very prestigious position
- You have very limited time, don't waste it

HOW YOUR WORK WILL  
CHANGE AFTER YOUR PHD

# FEELING RELIEVED?

- I'm done!
- No more overnight work, crazy deadlines...
- Now I don't have a boss anymore!

# NEW RESPONSIBILITIES

Teaching

Service

Fund raising

# TEACHING



# TEACHING

- Along with research, your primary job if you stay in academia
- Teaching priorities vary depending on the institutions

# HINTS

- Don't oversell yourself / overcommit
- Have fun and let your students have fun
- Connect your teaching with your research
- Don't teach what students can find on the Internet anyway

# SERVICE



# SERVICE TO YOUR INSTITUTION

- I agree this is by far the most boring part, but...
- Without it, excellence in education and research won't be possible

# SERVICE TO YOUR COMMUNITY

Reviewing

Conference organization

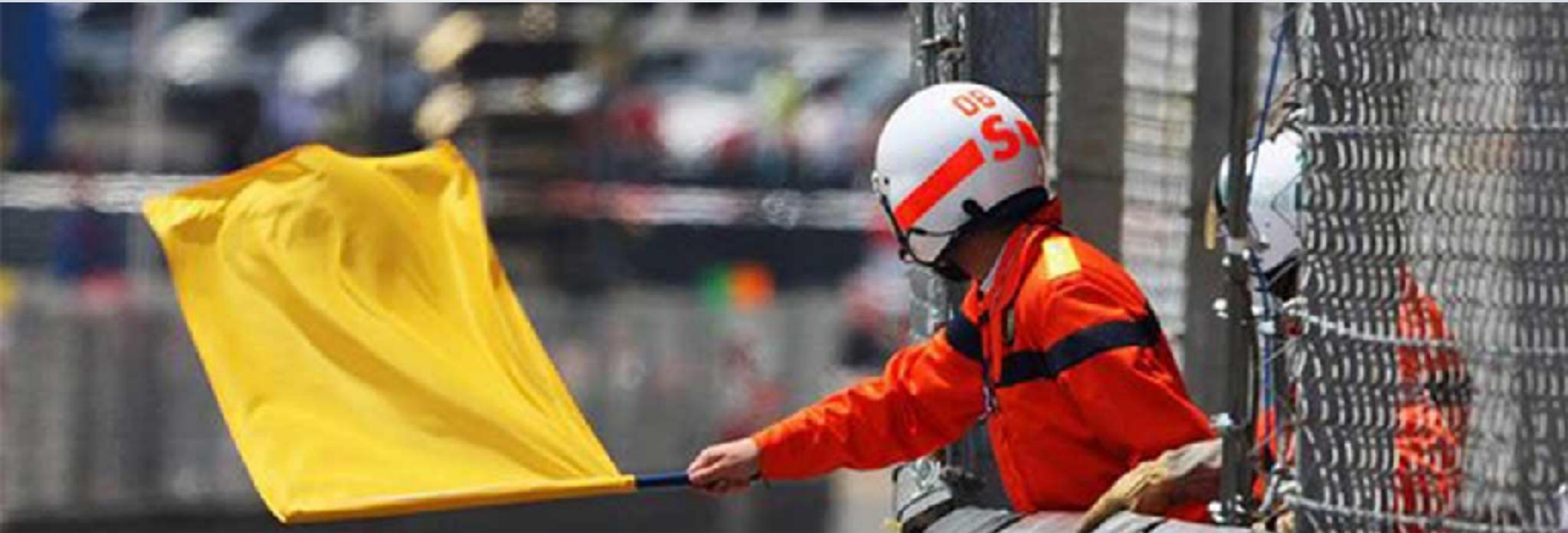
Other

# REVIEWING



- I admit I do a lot
- I care about my reviewing duties as much as I care about my research

# PROBLEMS



- Volume of submissions continuously increasing
- Not everybody available to reviewing

# HARD TO GET (GOOD) REVIEWERS

- Little credit, especially for journal reviewing
- Too much work in a very short time
- Conference reviewing might involve expensive PC meetings

CONFERENCES AND  
JOURNALS WON'T EXIST  
WITHOUT REVIEWERS!

# SO...

- If you write papers, you should also review
- Don't overcommit and feel free to say no, but not always
- Be timely / provide updates about your timeline

# WRITING GOOD REVIEWS

- Find reasons for acceptance, not for rejection
- **Don't write reviews you wouldn't like to receive**
- Explain why the paper should or should not be accepted
- Provide feedback to improve the paper

# FUND RAISING



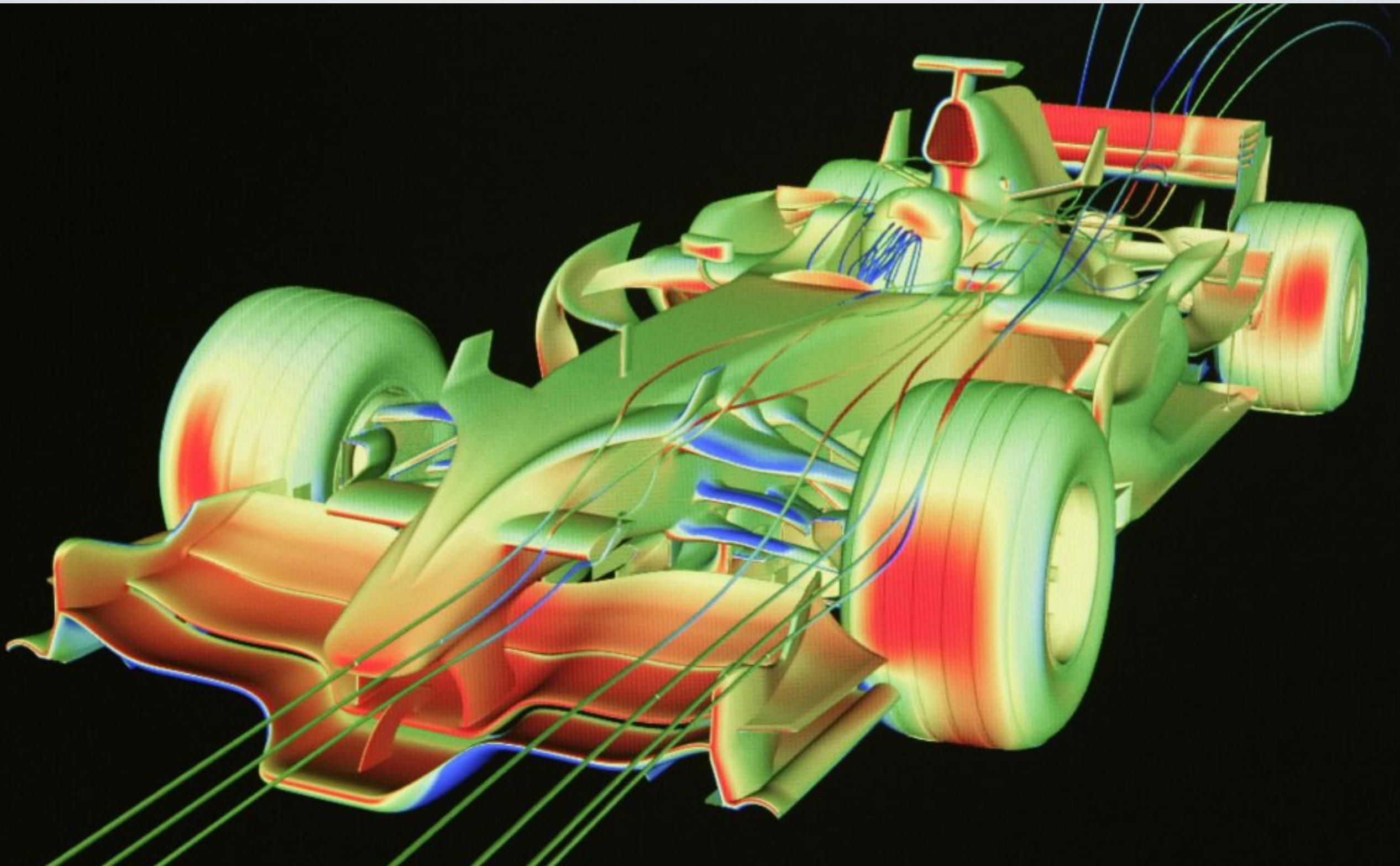
# FUND RAISING

- *Way too many models*
- You need to get familiar with the models of the country/area where you (will) work

# HINTS

- Fund raising functional to *your* research, not the other way around
- However, sometimes you need to fit your topics to grant calls

# WHAT ABOUT RESEARCH?



# WHAT ABOUT RESEARCH?

- Time to show your ability to develop your own path
- However, attacking a completely new topic takes time
- Make sure your research plan fits with your tenure timeline

# CHANGING ROLE



# CHANGING ROLE

You will get busier and busier pretty soon!

**Bad news: less time to do things you really like**

# MY PERSONAL SUGGESTION

Don't "abstract" yourself completely or you will loose contact with the reality

# ADVISING YOUR STUDENTS

- Always remember when you were a student
- Learn from your mistakes and from your advisor's mistakes

STILL SOMETIME LEFT FOR ME?



YES YOU NEED TO REFUEL!



# MANAGING LIFE

- Allocate time for your personal life
- Alternate periods in which you work under pressure with relatively calm period
- That's when you often have the best ideas!

# FINAL REMARKS



Contributions along different dimensions

Things can go wrong, keep it into account

Service is as important as research and teaching

Don't just become a manager

**Questions?**

[dipenta@unisannio.it](mailto:dipenta@unisannio.it), [@maxdipenta](https://twitter.com/maxdipenta)