IOWA STATE UNIVERSITY

# Empirical Research for Secure Software Development

Lotfi ben-Othmane

# Introduction

- Qualitative research is informed by qualitative data, such as explanation of people
- Quantitative research is concerned by quantitative data
- Data Science is data collection, data modeling and analysis, and decision making
- Empirical research is research using empirical evidence
  - Empirical evidence can be analyzed quantitatively or qualitatively

# Introduction

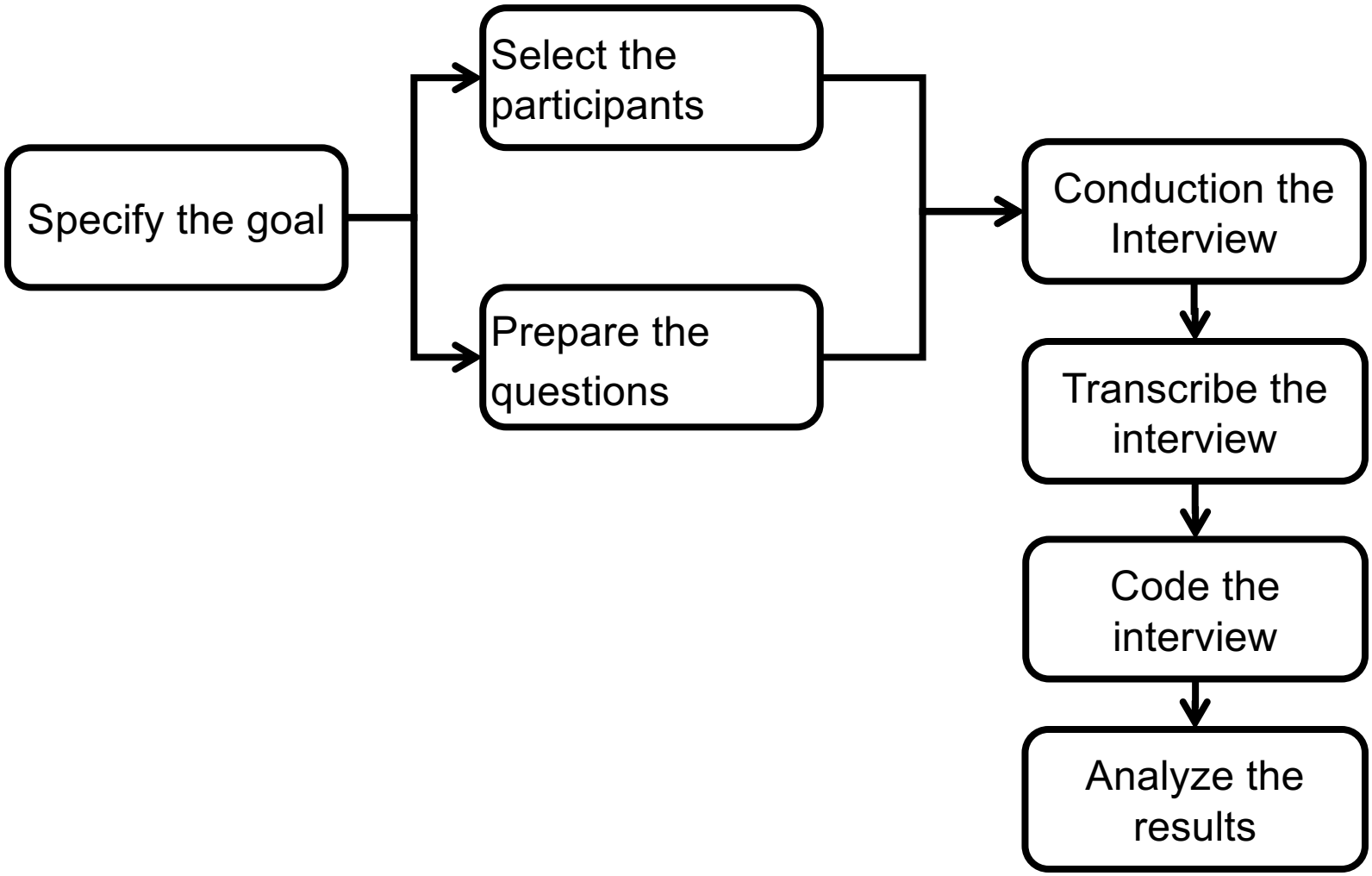
- 3 Methods

1. Qualitative analysis - Interviews
2. Quantitative analysis - Questionnaires
3. Data science- Regression

# Qualitative Research

1. What is it?
2. How did we use it?
3. What could you use it for?

# Research Method - Interviews



# Specify the Goal

Formulate clearly

1. What is the main research question?
2. What is the information type that you intend to get?

Example

- Q: What are the security practices in the organizations?
- R: A set of security practices that organization apply (like BISMM)

# Prepare the Questions

- Decide on question type
  - Open questions, closed questions, semi-open questions
- Formulate the interview protocol
  - Set of indirect questions that help to answer the main question
  - Advice: Use “how” and “why” question types and avoid “what” questions that require enumeration
- Test the interview protocol

# Select Participants

- Identify the groups of people that have the information you look for
- Better to select representatives from each group
  - Called maximum variation sampling
- Recall that Interviews are for exploratory research
  - Provide cumulative NOT comparative information

# Conduct the interview

- Interview opening
  - Collect information about the profile of interviewees
  - Disclose the interview goal and main research question
- Recording
  - Request permission for recording
  - Recording allows to fully capture the information
- Encourage interviewee to give details
  - Use prose such as “can you give an example”
- Interview closing
  - Ask for the possibility to get complementary information if needed

# Transcribe the Interview

- Write the interview in text that is easy to annotate and search
- Example of transcription tools: F4
- Transcriber should be familiar with the topic
  - Be careful about technical terms

# Code the Interview

Codes are themes and abstract concepts

- Example—code

  "I have been fixing security vulnerabilities for 5 years"

  into "experience in fixing vulnerabilities"

- Coding scheme allows grouping the codes into classes that together answer the main questions

- Coding is subjective

# Analyze the Results

- Develop a coding scheme
  - Identify common patterns in the interviews
- Extract data relevant to the research questions
- Use data frequency to get insights

# Threats to Validity Analysis

In General:

- Construct validity – Relation of the experiments to the hypothesis
- Internal validity - Causal relation of the experiments to the result
- Conclusion validity – Relation of the result to the experiments
- External validity - Condition to the generalization of the results

# Example of Qualitative Analysis

- Goal: Identify the factors that impact the vulnerabilities fixing time

- Result: The major factors that impact the fixing time

Factors Impacting the Effort Required to Fix Security Vulnerabilities

Lotfi ben Othmane, Golriz Chehrazi, Eric Bodden

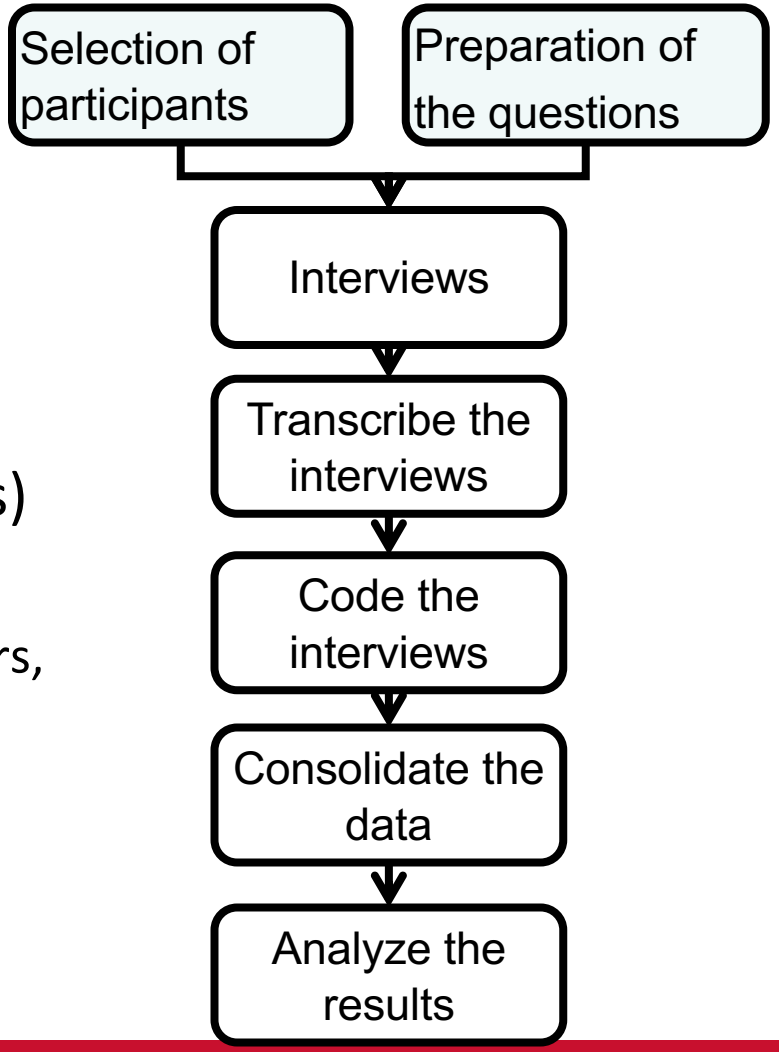
Petar Tsalovski, Achim Brucker, Philip Miseldine

(ISC 2015)

# The Study

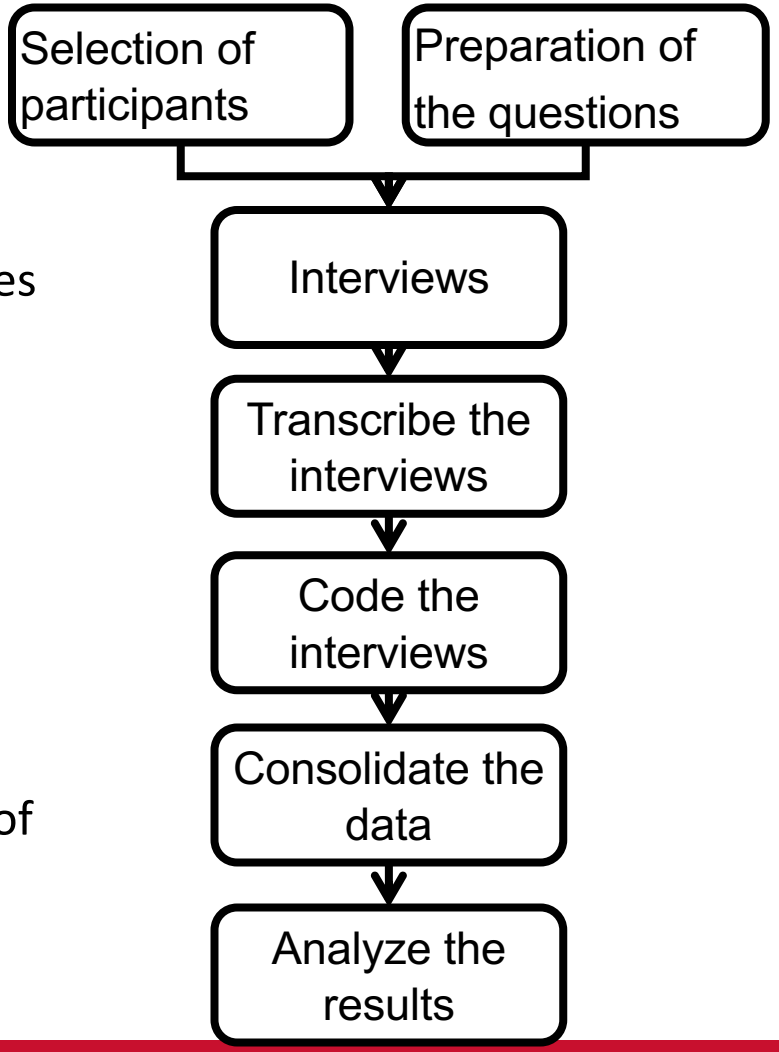
- The questions were designed to provide why fixing an example vulnerability took long time

- Number of participants 12 (12 hours)
  - 9 from Germany and 3 from India
  - Security experts, developers, coordinators, project leaders
  - NetWeaver experts, custom application experts, application experts



# The Study – Cont.

- Interviews were conducted from 8 to 12 Dec. 2014
- Each interview is transcribed into about 16 pages
- Identified 21 code classes from 3 sample interviews
- Coded each transcript in a report of 4 pages
- Each interviewee is asked to review the report of his interview



# The Study - cont.

- Coding examples
  - "Code injections are difficult to fix"

    ➔ vulnerability type => Vulnerability characteristics

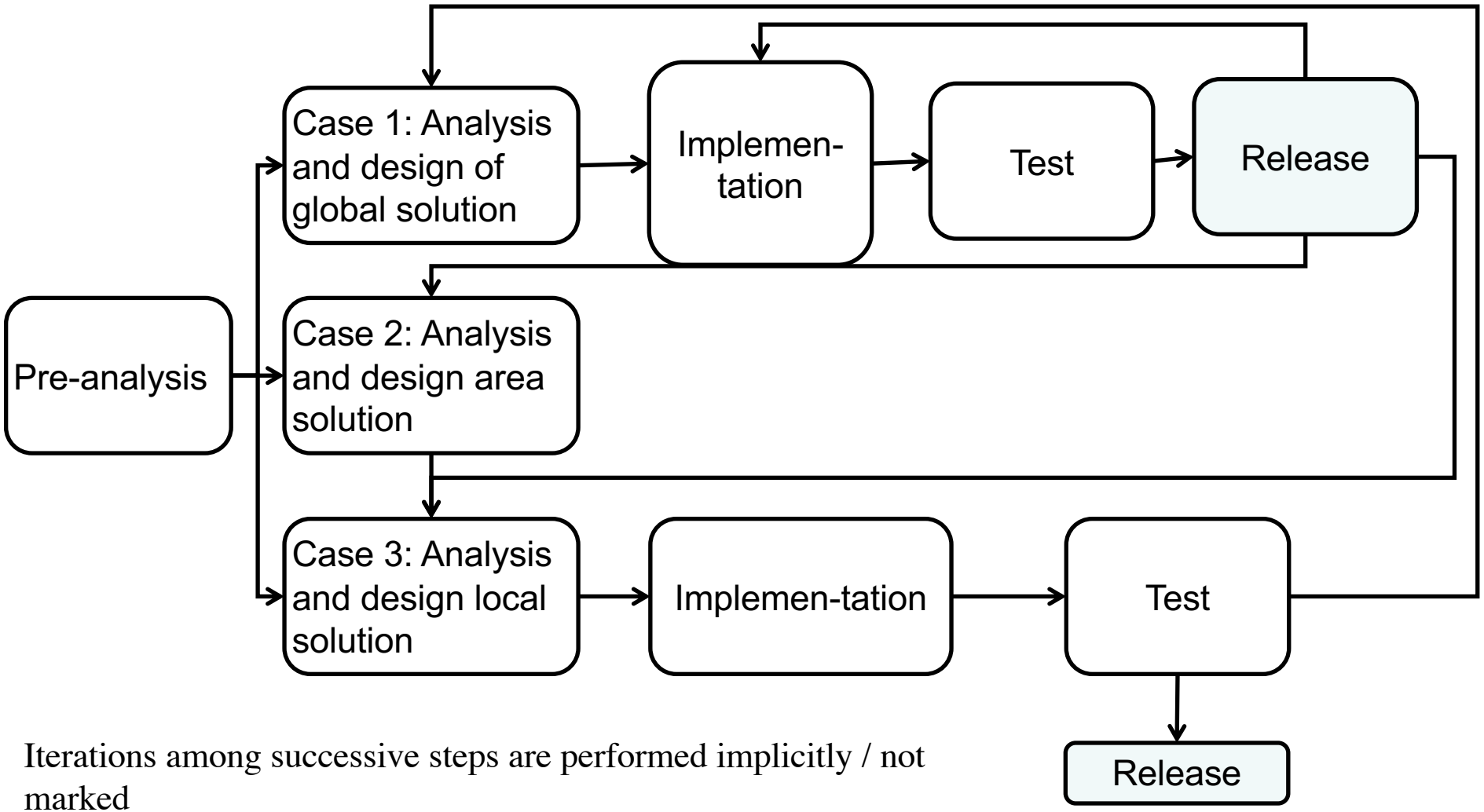
  - "If the function module is the same in all these 12 or 20 releases […] , then I just have to do one correction"

    ➔Similarity of code in the different releases => Software structure

# Factors that Impact the Vulnerability Fix time

| Factor categories | # of factors | Freq. |
|---|---|---|
| Vulnerabilities characteristics | 6 | 9 |
| Software structure | 19 | 10 |
| Technology diversification | 3 | 5 |
| Communication and collaboration | 7 | 8 |
| Availability and quality of information and documentation | 9 | 9 |
| Experience and knowledge | 12 | 11 |
| Code analysis tool | 4 | 4 |
| Other | 4 | 4 |

# Observed Fixing Process




Iterations among successive steps are performed implicitly / not marked

# Take-Away

- Vulnerability type is one among many factors (65) that impact the vulnerability fix time

- The 8 factor categories reflect the main areas for improving the vulnerability fixing processes
  - E.g., software structure, training, etc.

# Threats to Validity

- Control of the threats to the validity of the results
  - The interviewees are diversified
  - 2 researchers coded each interview and the results are consolidated
  - The participants validated the reports of their interviews
- Weaknesses
  - Used one method to identify the factors—interviewing experts
  - Interviewed only 2 developers
- External use
  - Diversity of product areas
  - Distribution of development teams

# Lessons learned

- The main interview questions shall help the interviewees to tell their own stories
  - "What" questions are inefficient to enumerate elements
- The participants sometimes have their own messages to deliver
- Vulnerability fixing processes are as many as the process participants

# Examples of Uses of the Method

- How do developers review code considering incremental software development?
- What are the software security practices that small/big organizations apply?

Thank you!

Questions?