# Git Helper System

Yihuan Dong
Niki Gitinabard
Linting Xue
Rui Zhi

Department of Computer Science
North Carolina State University

Spring 2016

NC STATE UNIVERSITY

# Outline

1. Motivation
2. Project Description
3. Solutions
   - Sol-1: Search Engine
   - Sol-2: Command Line Helper
   - Sol-3: Auto-reply Email System
4. Experiment
5. Result and Conclusion
6. Future Work

NC STATE UNIVERSITY

# Motivation

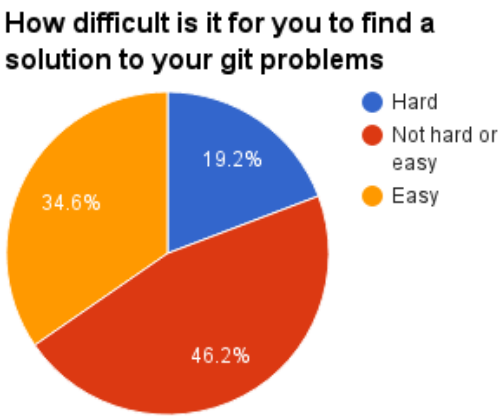



Figure: Expert User

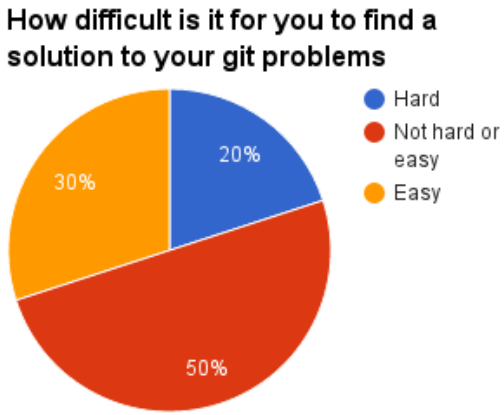



Figure: Novice User

# Outline

1. Motivation
2. Project Description
3. Solutions
   - Sol-1: Search Engine
   - Sol-2: Command Line Helper
   - Sol-3: Auto-reply Email System
4. Experiment
5. Result and Conclusion
6. Future Work

NC STATE UNIVERSITY

# Project Description

- Target User: Novice users who know a few of Git commands, don't know how to solve git errors properly

# Project Description

- Target User: Novice users who know a few of Git commands, don't know how to solve git errors properly
- Goal: Design tools to help people use Git comfortably

NC STATE UNIVERSITY

# Outline

1. Motivation
2. Project Description
3. Solutions
   - Sol-1: Search Engine
   - Sol-2: Command Line Helper
   - Sol-3: Auto-reply Email System
4. Experiment
5. Result and Conclusion
6. Future Work

NC STATE UNIVERSITY

# Solution

- Solution 1 - Git Helper Search Engine

# Solution

- Solution 1 - Git Helper Search Engine
- Solution 2 - Command Line Helper

# Solution

- Solution 1 - Git Helper Search Engine
- Solution 2 - Command Line Helper
- Solution 3 - Auto-replay Email System

# Outline

1 Motivation

2 Project Description

3 Solutions
- Sol-1: Search Engine
- Sol-2: Command Line Helper
- Sol-3: Auto-reply Email System

4 Experiment

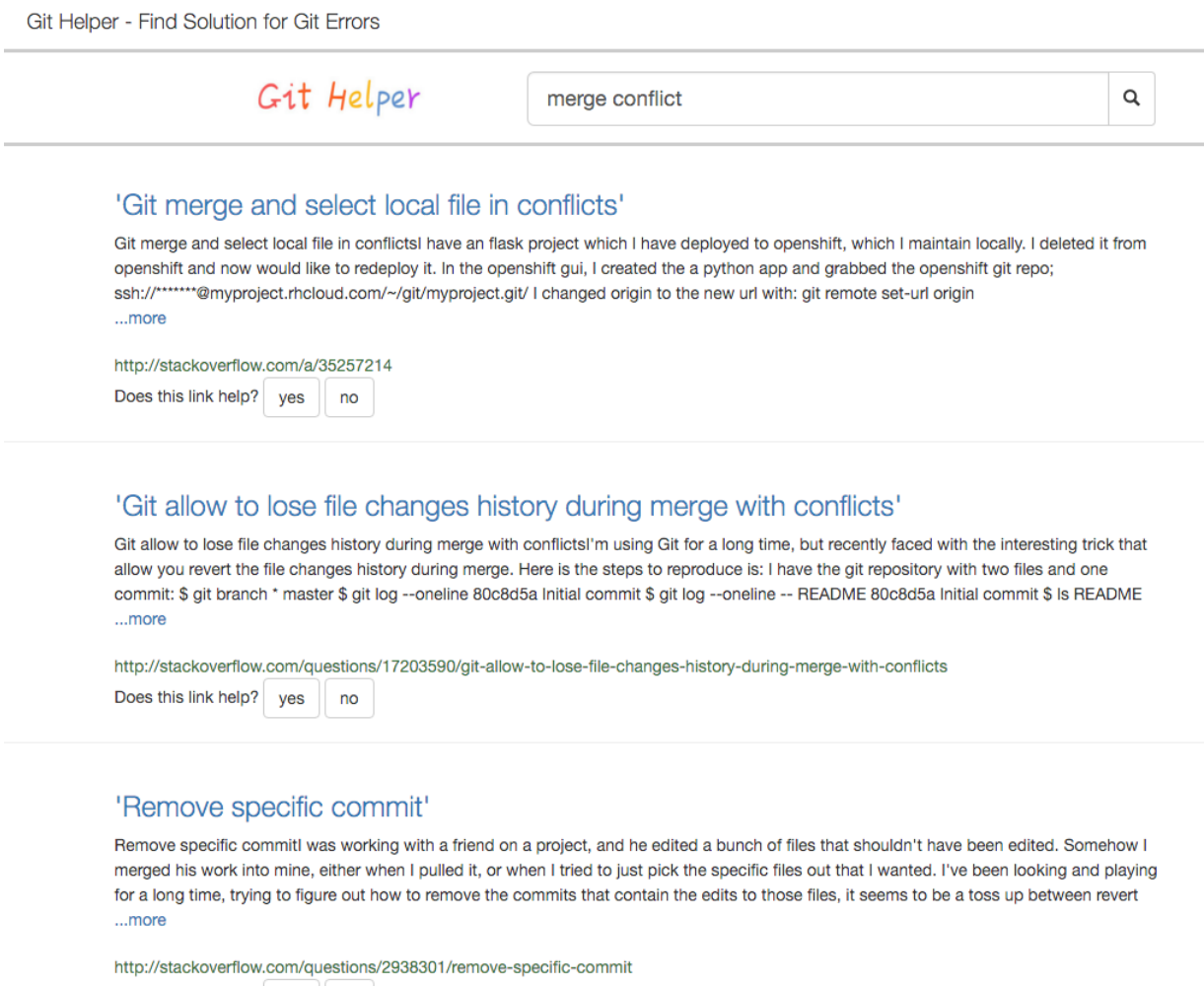
5 Result and Conclusion

6 Future Work

# Sol-1: Search Engine

- Design Process
- Implementation Details
  - Webpage Crawler
  - Google App Engine Documentation and Indexes

- Advantages:
  - Convenience
  - Generalizability
  - Multiple Solutions
- Disadvantages:
  - No Fuzzy Search
  - Solution Quality Varies




//youtu.be/CupW8dWM91Y

NC STATE UNIVERSITY

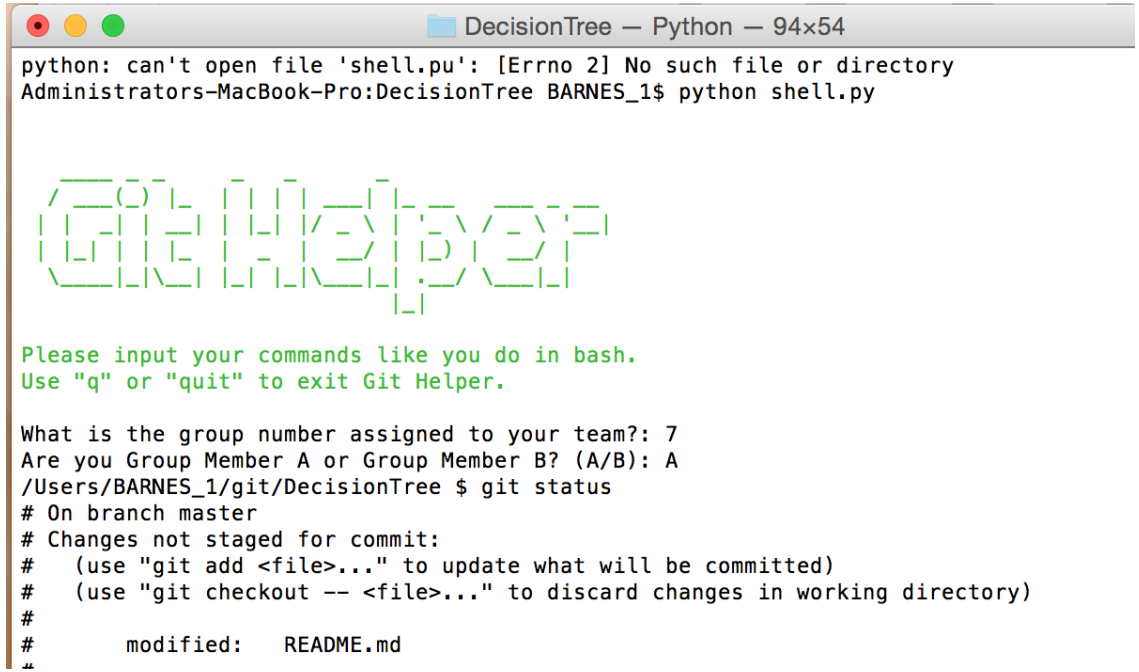
# Outline


1. Motivation
2. Project Description
3. Solutions
   - Sol-1: Search Engine
   - Sol-2: Command Line Helper
   - Sol-3: Auto-reply Email System
4. Experiment
5. Result and Conclusion
6. Future Work


NC STATE UNIVERSITY

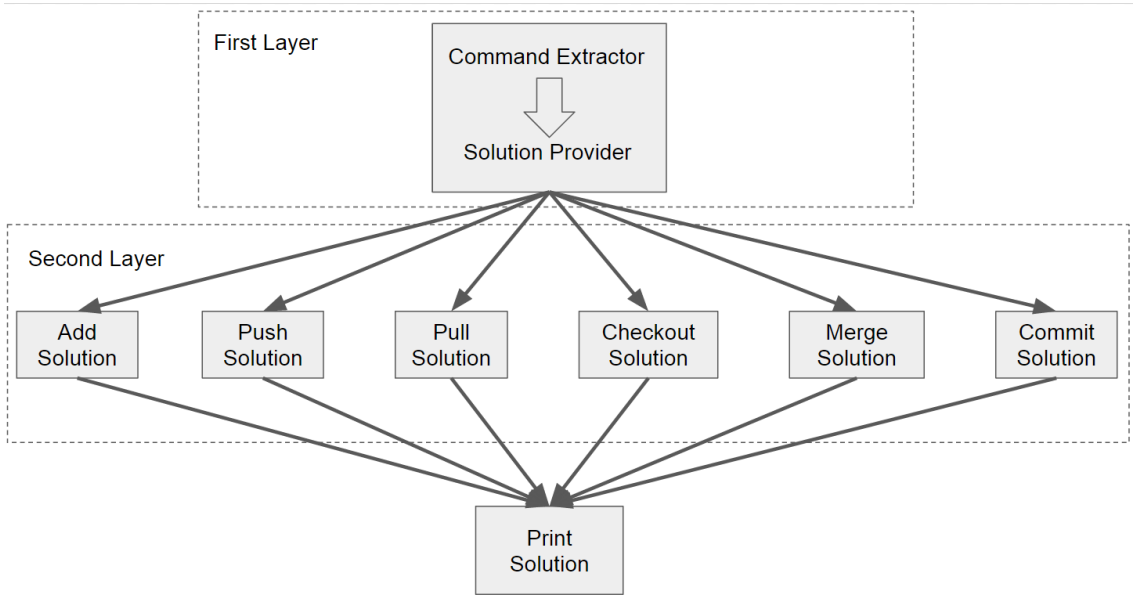
# Sol-3: Command Line Helper

- Customized Shell
- Error message and solution collection



# Sol-3: Command Line Helper

- Decision Tree
- Advantages
  - Time Saving
  - Accurate
- Disadvantages
  - Solution Coverage
  - Expert Knowledge
- `https://youtu.be/3DvWSL3XM5w`




NC STATE UNIVERSITY

# Outline

1. Motivation
2. Project Description
3. Solutions
   - Sol-1: Search Engine
   - Sol-2: Command Line Helper
   - Sol-3: Auto-reply Email System
4. Experiment
5. Result and Conclusion
6. Future Work

NC STATE UNIVERSITY

# Creating the Database

- Crawl web contents

# Creating the Database

- Crawl web contents
- Tokenize crawled contents
  - Remove punctuations, Unicode characters, stop words
  - Change capitals to lowercase
  - $\Rightarrow$ Obtain 14480 unique words

# Creating the Database

- Crawl web contents
- Tokenize crawled contents
  - Remove punctuations, Unicode characters, stop words
  - Change capitals to lowercase
  - $\Rightarrow$ Obtain 14480 unique words
- Obtain TF-IDF matrix (1311 * 14480)

# Creating the Database

- Crawl web contents
- Tokenize crawled contents
  - Remove punctuations, Unicode characters, stop words
  - Change capitals to lowercase
  - $\Rightarrow$ Obtain 14480 unique words
- Obtain TF-IDF matrix (1311 * 14480)
- Save TF-IDF matrix into CSV file as database for query

# Mail Server

- Checking inbox for git_helper@yahoo.com every 20 sec

# Mail Server

- Checking inbox for git_helper@yahoo.com every 20 sec
- For any new message, check the database for answer

# Mail Server

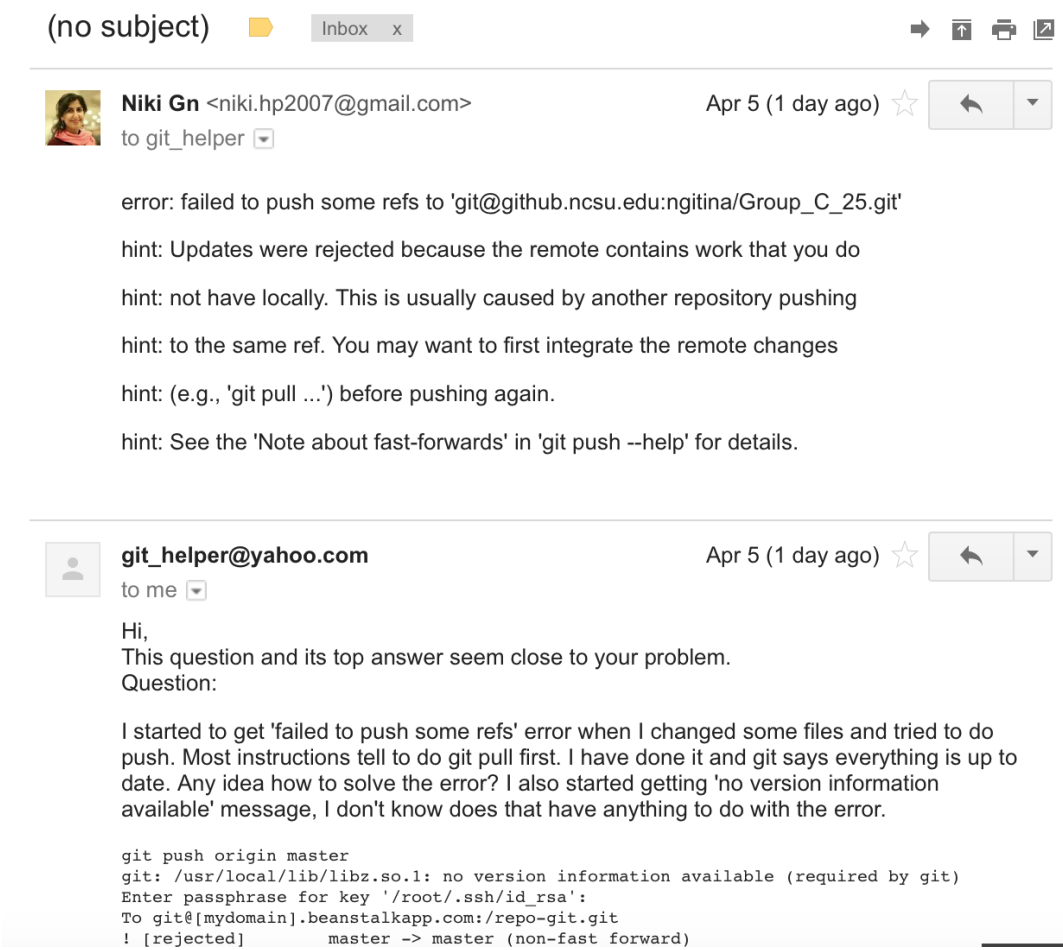
- Checking inbox for git_helper@yahoo.com every 20 sec
- For any new message, check the database for answer
- Reply with the found answer

# Mail Server

- Checking inbox for git_helper@yahoo.com every 20 sec
- For any new message, check the database for answer
- Reply with the found answer
- Record the email as answered

# Mail Server

- Checking inbox for git_helper@yahoo.com every 20 sec
- For any new message, check the database for answer
- Reply with the found answer
- Record the email as answered
- Demo



(no subject) Inbox x

**Niki Gn** <niki.hp2007@gmail.com> Apr 5 (1 day ago)
to git_helper

error: failed to push some refs to 'git@github.ncsu.edu:ngitina/Group_C_25.git'

hint: Updates were rejected because the remote contains work that you do

hint: not have locally. This is usually caused by another repository pushing

hint: to the same ref. You may want to first integrate the remote changes

hint: (e.g., 'git pull ...') before pushing again.

hint: See the 'Note about fast-forwards' in 'git push --help' for details.

**git_helper@yahoo.com** Apr 5 (1 day ago)
to me

Hi,
This question and its top answer seem close to your problem.
Question:

I started to get 'failed to push some refs' error when I changed some files and tried to do push. Most instructions tell to do git pull first. I have done it and git says everything is up to date. Any idea how to solve the error? I also started getting 'no version information available' message, I don't know does that have anything to do with the error.

```
git push origin master
git: /usr/local/lib/libz.so.1: no version information available (required by git)
Enter passphrase for key '/root/.ssh/id_rsa':
To git@[mydomain].beanstalkapp.com:/repo-git.git
! [rejected]        master -> master (non-fast forward)
```

NC STATE UNIVERSITY

# Query and Find the Answer

- Tokenize message to get key words

# Query and Find the Answer

- Tokenize message to get key words
- Sum up Tf-Idf values of key words for each entry in database

# Query and Find the Answer

- Tokenize message to get key words
- Sum up Tf-Idf values of key words for each entry in database
- Rank database pages

# Query and Find the Answer

- Tokenize message to get key words
- Sum up Tf-Idf values of key words for each entry in database
- Rank database pages
- Return the top post

# Advantages and Disadvantages

- Advantages
  - No need to install anything for user
  - No need to find the best answer for user
  - Ignore the personal and non-relevant part of the error
  - No need to expert user information

# Advantages and Disadvantages

- Advantages
  - No need to install anything for user
  - No need to find the best answer for user
  - Ignore the personal and non-relevant part of the error
  - No need to expert user information
- Disadvantages
  - The first question might not be actually useful
  - They need to wait for email to be answered

# Outline

1. Motivation
2. Project Description
3. Solutions
   - Sol-1: Search Engine
   - Sol-2: Command Line Helper
   - Sol-3: Auto-reply Email System
4. Experiment
5. Result and Conclusion
6. Future Work

NC STATE UNIVERSITY

# Properties

- 56 Students

NC STATE UNIVERSITY

# Properties

- 56 Students
- Undergraduate, Having CSC216 (Programming Concepts-Java)

NC STATE UNIVERSITY

# Properties

- 56 Students
- Undergraduate, Having CSC216 (Programming Concepts-Java)
- Baseline: Use any tool you want

# Properties

- 56 Students
- Undergraduate, Having CSC216 (Programming Concepts-Java)
- Baseline: Use any tool you want
- 4 groups for each tool and baseline
  - Based on git proficiency survey
  - Grouped one week before experiment
- Teams of 2 people working together in class
  - Labeled as A and B
  - Assigned to a git repository to work on
  - Assigned stepwise procedure to face a problem and solve it by their tool

NC STATE UNIVERSITY

# Logging

- All solutions use a shell for time recording
- All the commands and errors entered in shell are recorded

# Participation

From 56 participants:

- 45 post surveys completed
- 22 useful log files gathered

# Outline


1. Motivation
2. Project Description
3. Solutions
   - Sol-1: Search Engine
   - Sol-2: Command Line Helper
   - Sol-3: Auto-reply Email System
4. Experiment
5. Result and Conclusion
6. Future Work


NC STATE UNIVERSITY

# Experiment Statistics

| Solutions | Participant Users | Useful Log Files | Satisfaction Surveys |
|---|---|---|---|
| Search Engine | 14 | **6** | **10** |
| Email Server | 14 | **4** | **13** |
| Decision Trees | 14 | **5** | **12** |
| Baseline | 14 | **7** | 10 |
| Total | 56 | 22 | 45(35) |

NC STATE UNIVERSITY

# Interaction Log Files Analysis

| Solutions | Average Spending Time | |
|---|---|---|
| | Complete Test | Solve Error |
| **Search Engine** | $17'54''$ | $5'56''$ |
| **Email Server** | $11'30''$ | $2'45''$ |
| **Command Line** | $\mathbf{11'24''}$ | $\mathbf{1'49''}$ |
| **Baseline** | $17'22''$ | $5'42''$ |

NC STATE UNIVERSITY

# Interaction Log Files Analysis

| Solutions | Average Spending Time | |
|---|---|---|
| | Complete Test | Solve Error |
| **Search Engine** | $17'54''$ | $5'56''$ |
| **Email Server** | $11'30''$ | $2'45''$ |
| **Command Line** | $\mathbf{11'24''}$ | $\mathbf{1'49''}$ |
| **Baseline** | $17'22''$ | $5'42''$ |

$\Rightarrow$Command line is the most efficient

NC STATE UNIVERSITY

# Satisfaction Survey Analysis

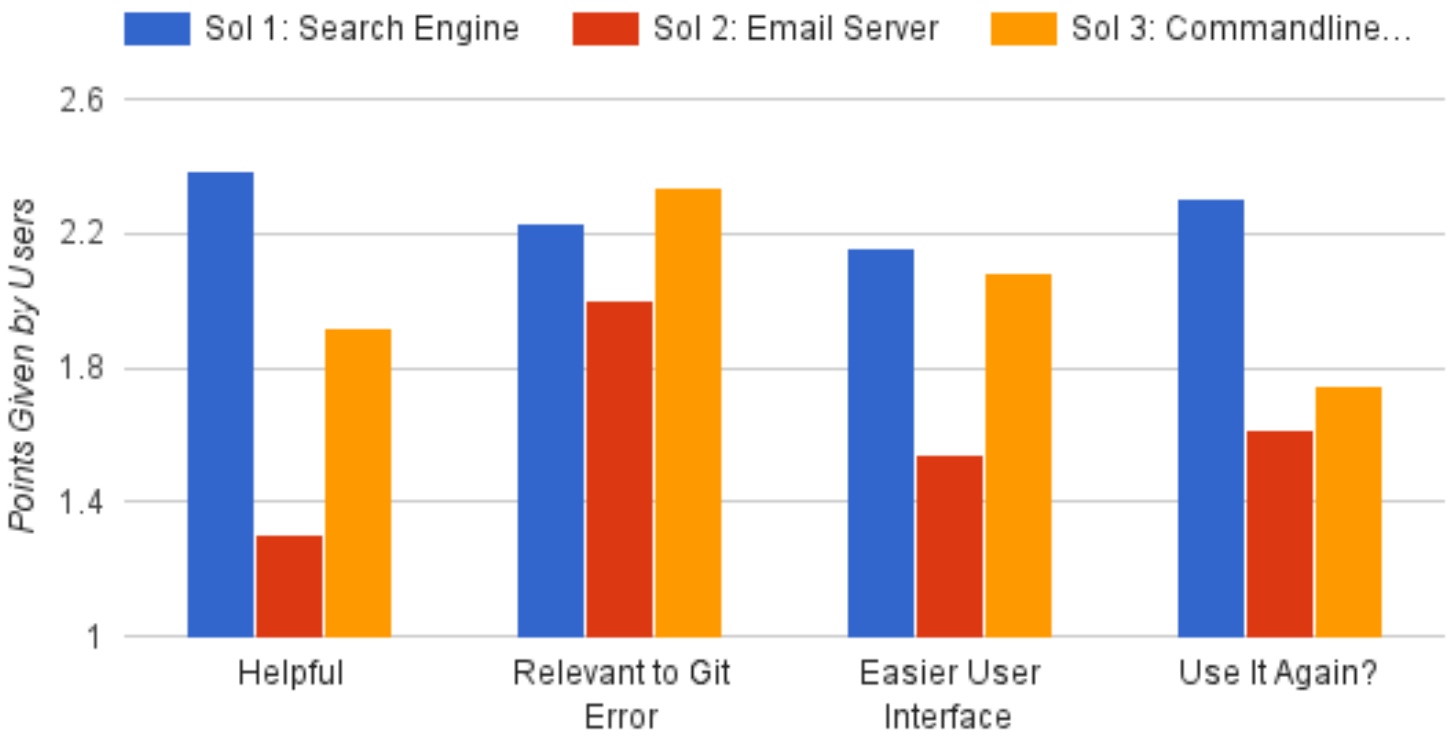



Figure: Satisfaction Survey Comparison Between Three Solutions

NC STATE UNIVERSITY

# Conclusion

- Best Solution: **Command Line helper**

# Conclusion

- Best Solution: **Command Line helper**
- Reasons:
  - Most efficient tool
  - Most relevant solution

# Outline

1. Motivation
2. Project Description
3. Solutions
   - Sol-1: Search Engine
   - Sol-2: Command Line Helper
   - Sol-3: Auto-reply Email System
4. Experiment
5. Result and Conclusion
6. Future Work

NC STATE UNIVERSITY

# Future Work

- Search Engine:
  - Expand database
  - Support fuzzy search

# Future Work

- Search Engine:
  - Expand database
  - Support fuzzy search
- Auto-reply Email System:
  - Support new features, e.g. next best result
  - Expand email format support

# Future Work

- Search Engine:
  - Expand database
  - Support fuzzy search
- Auto-reply Email System:
  - Support new features, e.g. next best result
  - Expand email format support
- Command Line Helper:
  - Add new layers to check repository status
  - Find ways for easy expansion and collaboration

# Questions

NC STATE UNIVERSITY