CS8395 Advanced Topics in Software Engineering

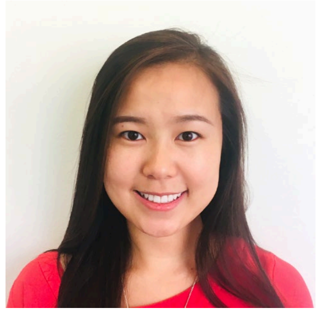
Yu Huang

Vanderbilt University

yu.huang@vanderbilt.edu

Course Information

- **Instructor**: Yu Huang
    - yu.huang@vanderbilt.edu
    - PhD in Computer Science, UMich
    - MS in Computer Engineering, UVa
    - Software Engineering: human factors & human-centered AI4SE, OSS, CS Education
    - https://yuhuang-lab.github.io/



- **Course meetings**
    - ESB 048
    - TR 1:15pm – 2:30pm
- **Office hours**
    - Tuesday 3:00pm – 4:00pm, Sony 4108

Course Information

- **GSI:** Yueke Zhang
  - Yueke.zhang@vanderbilt.edu
- **Office hours**
  - Thursday 9:00am – 10:00am, Sony 4118

What is the worst thing that can happen if you say something wrong in this class?

What is the worst thing that can happen if you say something wrong in this class?

Who are you?

- Introduction
    - Who are you
    - What is your research direction?
        - If you have been working on a research project, what is that?
        - If you haven’t, what are you interested in doing in the future?
        - “I don’t know” → totally fine!
    - A fun fact about you

Administrative Information

- Course website: **https://huang.isis.vanderbilt.edu/cs8395**

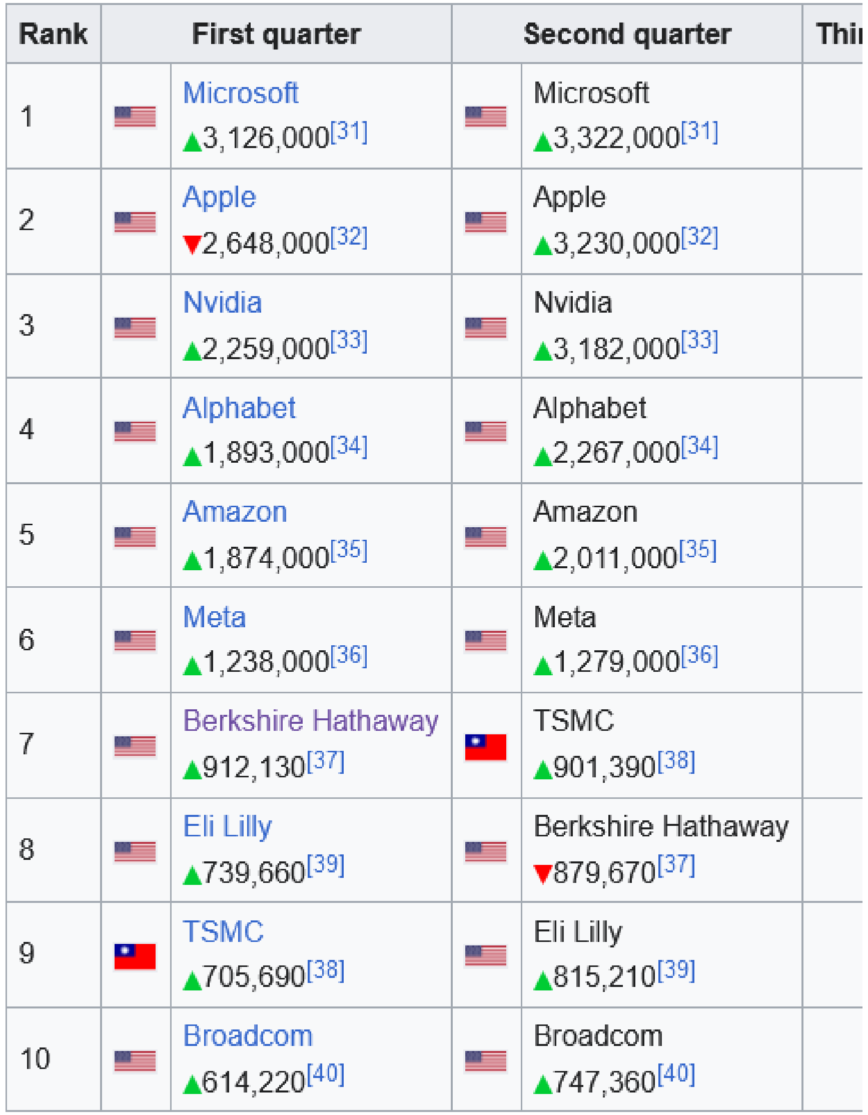
  You can find everything on the course website (soon to be finalized)!

- Piazza
  - Announcement, Q&A, discussion, etc.

- Brightspace
  - Only for written assignment submission and grade release

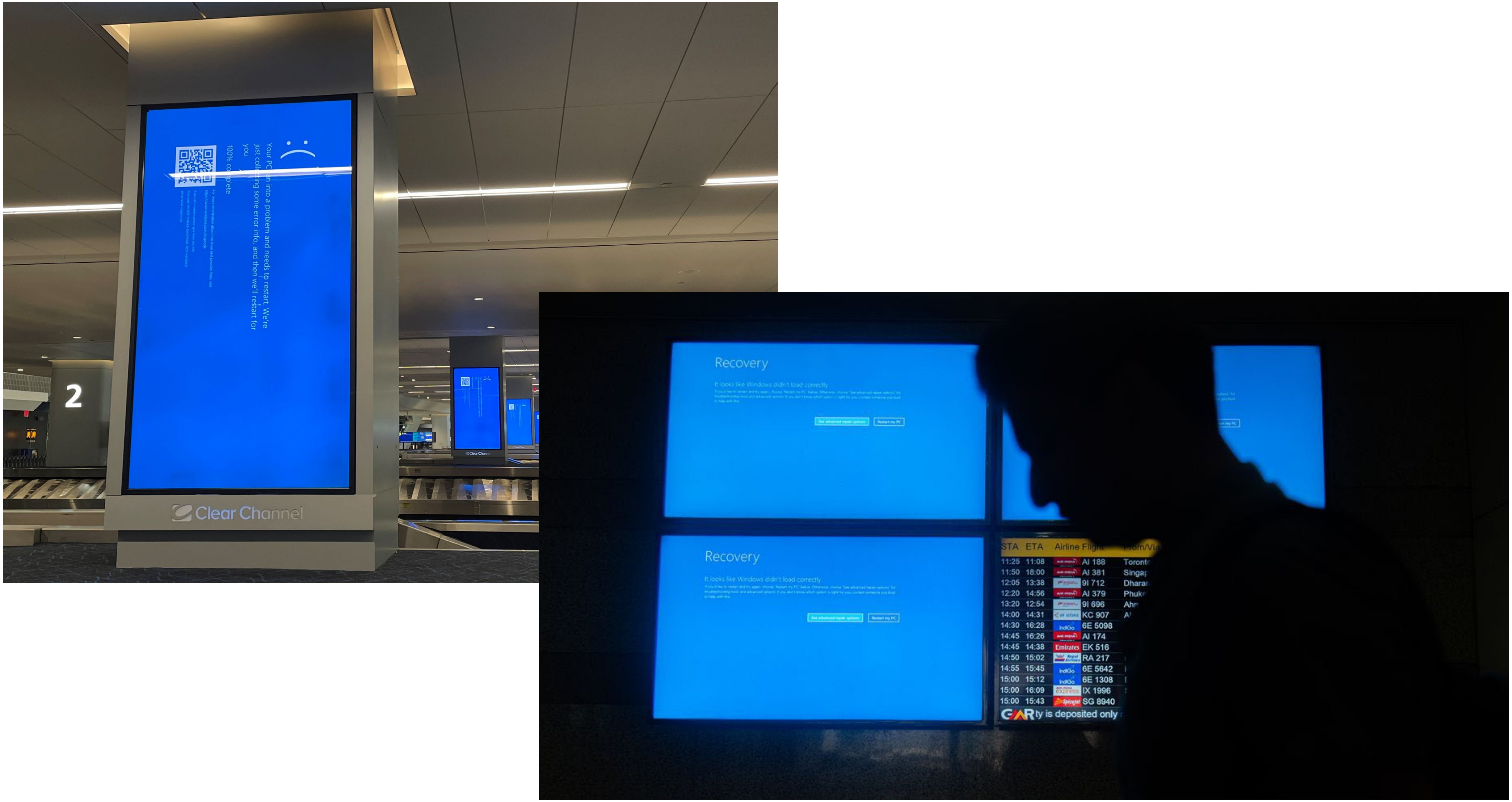
Why should you study Software Engineering?

List of public corporations by market cap - 2024



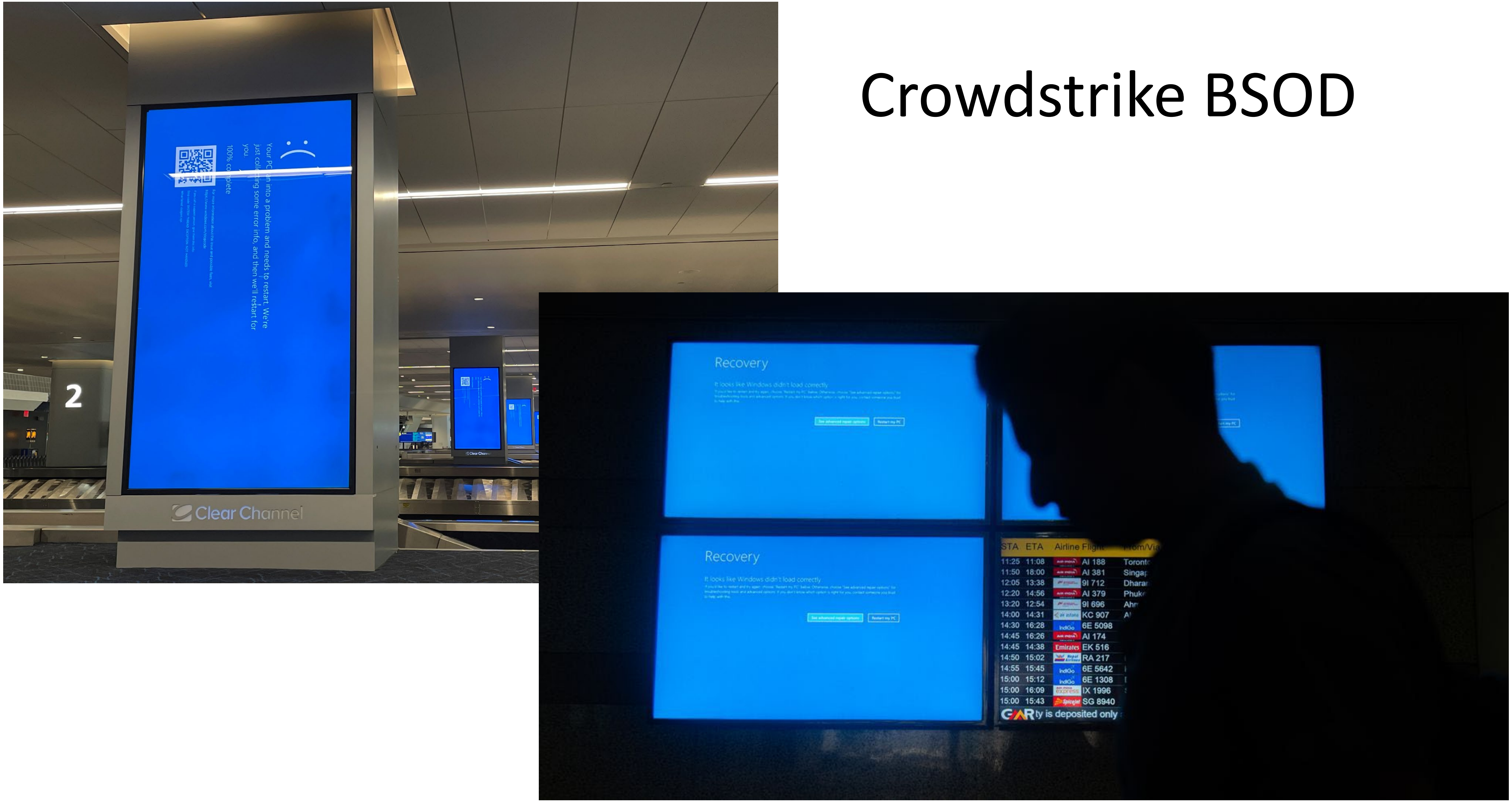
Rank	First quarter		Second quarter		Thi
1	🇺🇸	Microsoft ▲3,126,000[31]	🇺🇸	Microsoft ▲3,322,000[31]	
2	🇺🇸	Apple ▼2,648,000[32]	🇺🇸	Apple ▲3,230,000[32]	
3	🇺🇸	Nvidia ▲2,259,000[33]	🇺🇸	Nvidia ▲3,182,000[33]	
4	🇺🇸	Alphabet ▲1,893,000[34]	🇺🇸	Alphabet ▲2,267,000[34]	
5	🇺🇸	Amazon ▲1,874,000[35]	🇺🇸	Amazon ▲2,011,000[35]	
6	🇺🇸	Meta ▲1,238,000[36]	🇺🇸	Meta ▲1,279,000[36]	
7	🇺🇸	Berkshire Hathaway ▲912,130[37]	🇹🇼	TSMC ▲901,390[38]	
8	🇺🇸	Eli Lilly ▲739,660[39]	🇺🇸	Berkshire Hathaway ▼879,670[37]	
9	🇹🇼	TSMC ▲705,690[38]	🇺🇸	Eli Lilly ▲815,210[39]	
10	🇺🇸	Broadcom ▲614,220[40]	🇺🇸	Broadcom ▲747,360[40]	

Source: https://en.wikipedia.org/wiki/List_of_public_corporations_by_market_capitalization (as of June 30, 2024)

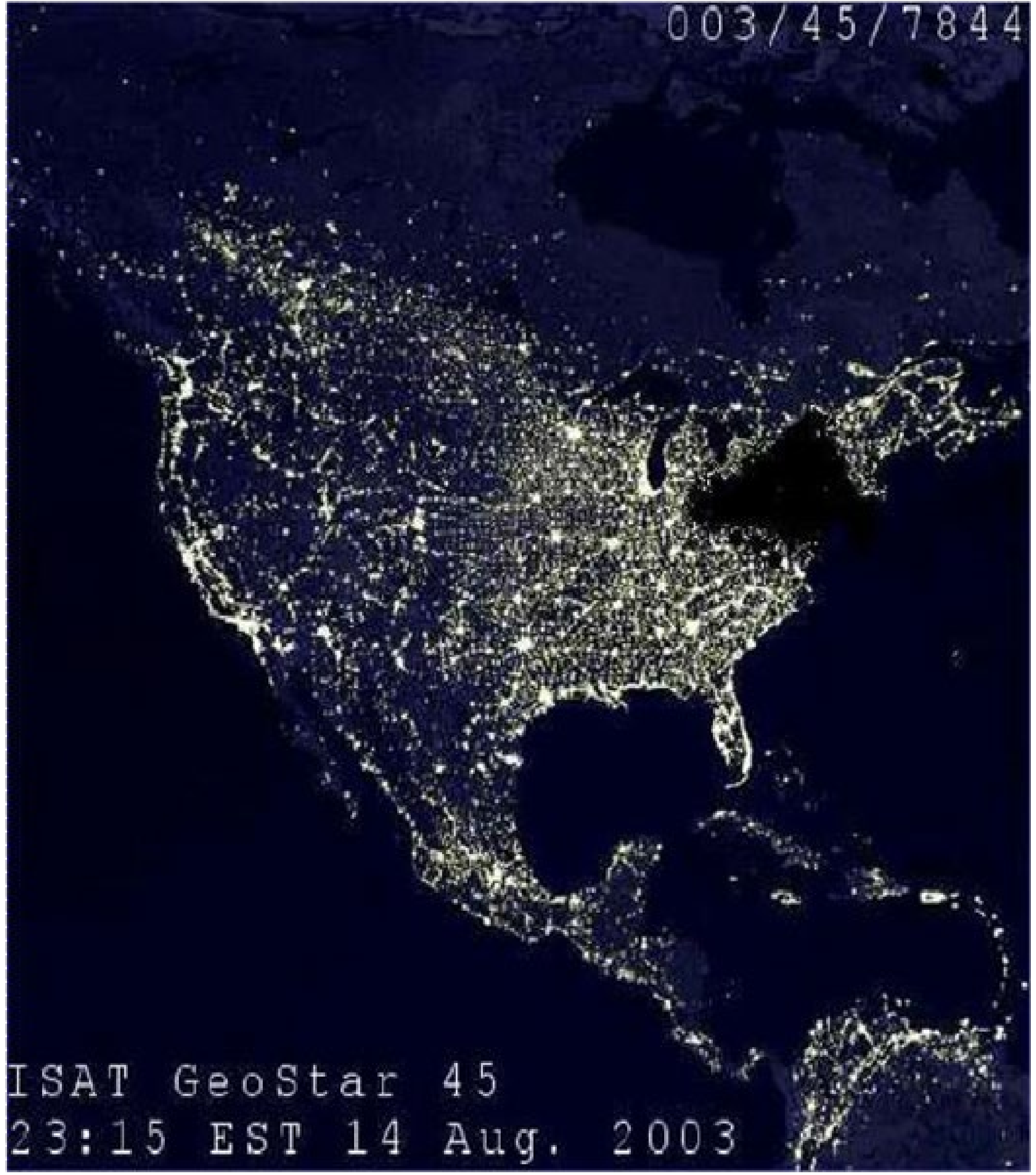


:(
Your PC ran into a problem and needs to restart. We're just collecting some error info, and then we'll restart for you.
100% complete
2
Clear Channel
Recovery
It looks like Windows didn't load correctly
Recovery
It looks like Windows didn't load correctly
STA ETA Airline Flight From/Via
11:25 11:08 AI 188 Toronto
11:50 18:00 AI 381 Singap
12:05 13:38 9I 712 Dharan
12:20 14:56 AI 379 Phuke
13:20 12:54 9I 696 Ahm
14:00 14:31 KC 907
14:30 16:28 6E 5098
14:45 16:26 AI 174
14:45 14:38 Emirates EK 516
14:50 15:02 RA 217
14:55 15:45 IndiGo 6E 5642
15:00 15:12 IndiGo 6E 1308
15:00 16:09 IX 1996
15:00 15:43 SpiceJet SG 8940
GMR ty is deposited only

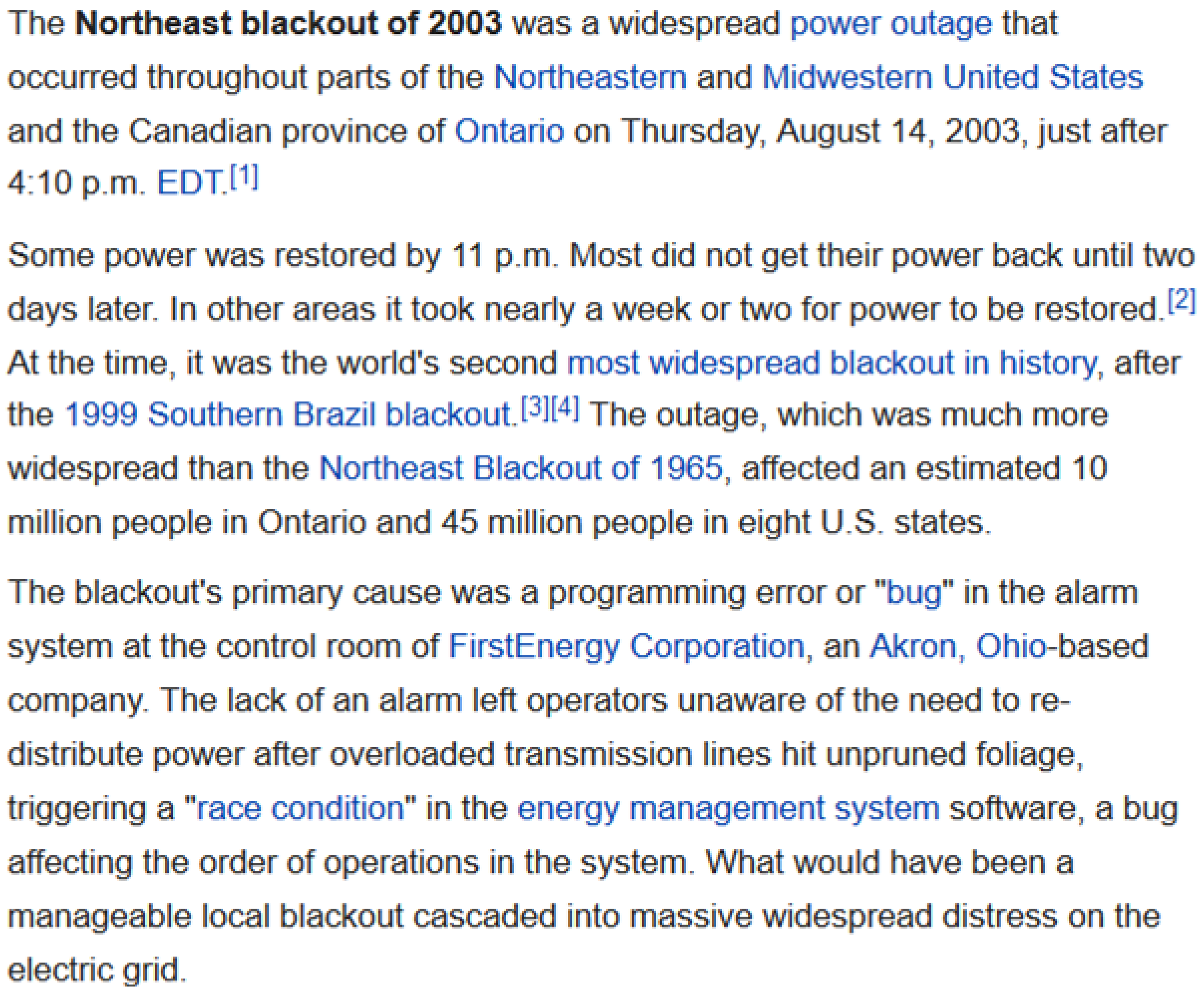
Crowdstrike BSOD


:(
Your PC ran into a problem and needs to restart. We're just collecting some error info, and then we'll restart for you.
100% complete
2
Clear Channel
Recovery
It looks like Windows didn't load correctly
See advanced repair options
Restart my PC
Recovery
It looks like Windows didn't load correctly
See advanced repair options
Restart my PC
STA ETA Airline Flight From/Via
11:25 11:08 AI 188 Toronto
11:50 18:00 AI 381 Singap
12:05 13:38 9I 712 Dharam
12:20 14:56 AI 379 Phuke
13:20 12:54 9I 696 Ahm
14:00 14:31 KC 907
14:30 16:28 6E 5098
14:45 16:26 AI 174
14:45 14:38 Emirates EK 516
14:50 15:02 RA 217
14:55 15:45 IndiGo 6E 5642
15:00 15:12 IndiGo 6E 1308
15:00 16:09 IX 1996
15:00 15:43 SpiceJet SG 8940
GMR ty is deposited only

What is the wrong with this picture?



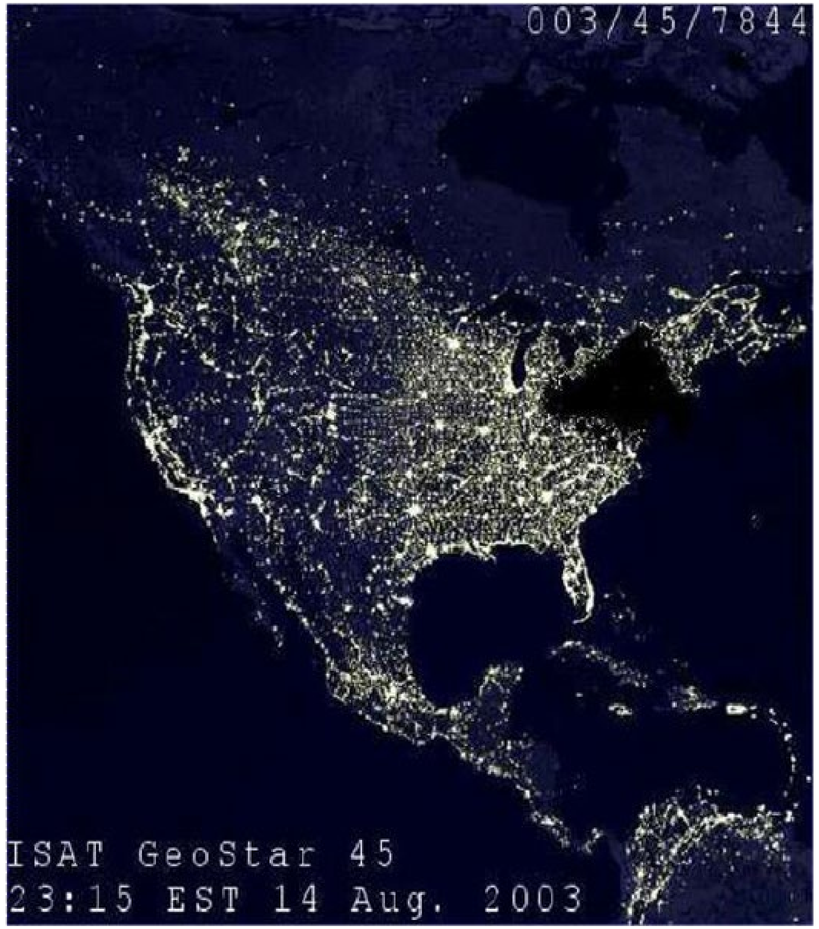
Software is Critical: Power



The **Northeast blackout of 2003** was a widespread power outage that occurred throughout parts of the Northeastern and Midwestern United States and the Canadian province of Ontario on Thursday, August 14, 2003, just after 4:10 p.m. EDT.[1]

Some power was restored by 11 p.m. Most did not get their power back until two days later. In other areas it took nearly a week or two for power to be restored.[2] At the time, it was the world's second most widespread blackout in history, after the 1999 Southern Brazil blackout.[3][4] The outage, which was much more widespread than the Northeast Blackout of 1965, affected an estimated 10 million people in Ontario and 45 million people in eight U.S. states.

The blackout's primary cause was a programming error or "bug" in the alarm system at the control room of FirstEnergy Corporation, an Akron, Ohio-based company. The lack of an alarm left operators unaware of the need to re-distribute power after overloaded transmission lines hit unpruned foliage, triggering a "race condition" in the energy management system software, a bug affecting the order of operations in the system. What would have been a manageable local blackout cascaded into massive widespread distress on the electric grid.

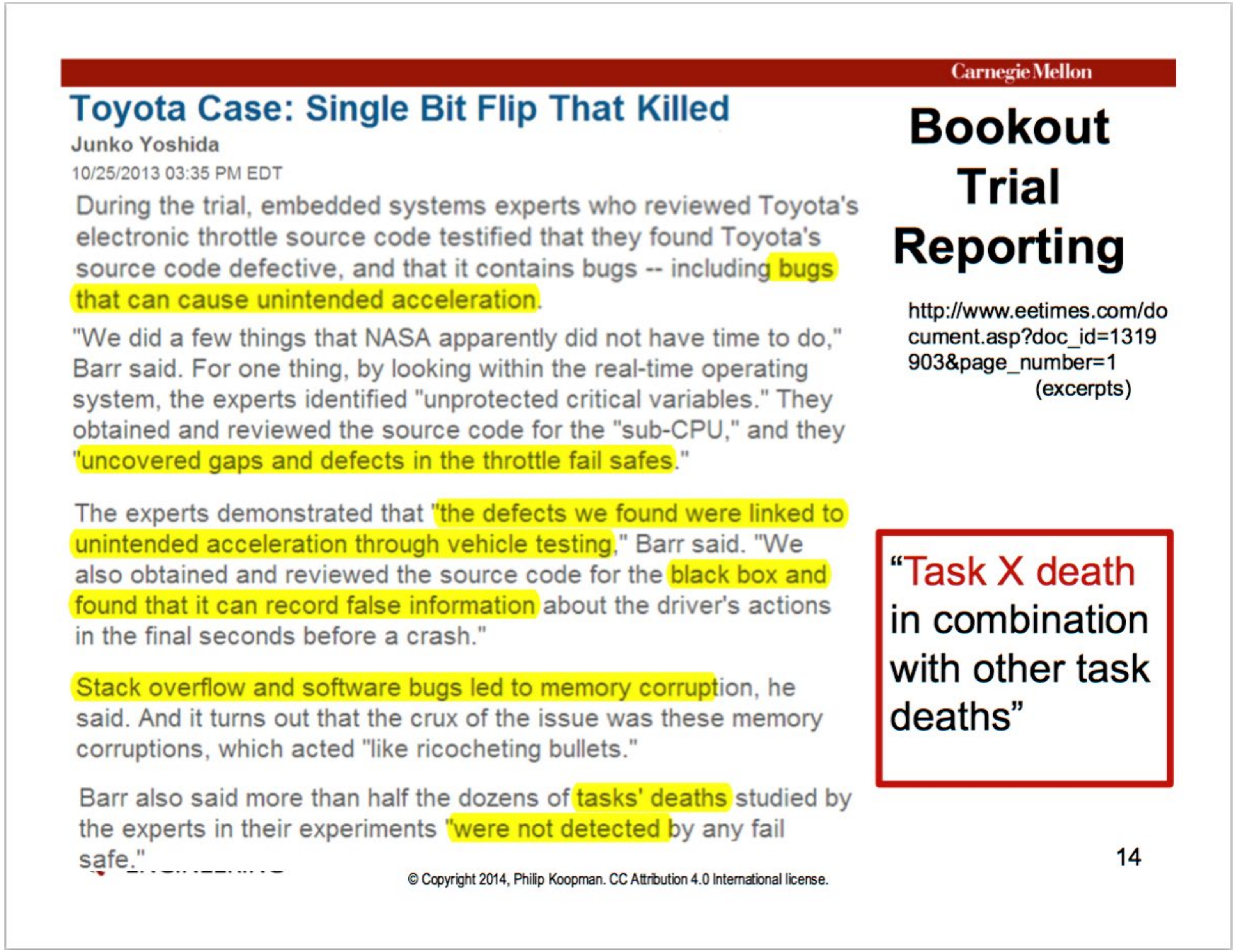


Software is Critical: Defense

- Quoting an Air Force lieutenant general, “The only thing you can do with an F-22 that does not require software is take a picture of it.”

[Crouching Dragon, Hidden Software: Software in DOD Weapon Systems (Ferguson, IEEE Software, 2001)]

Software is Critical: Driving



Carnegie Mellon

Toyota Case: Single Bit Flip That Killed

Junko Yoshida

10/25/2013 03:35 PM EDT

During the trial, embedded systems experts who reviewed Toyota's electronic throttle source code testified that they found Toyota's source code defective, and that it contains bugs -- including bugs that can cause unintended acceleration.

"We did a few things that NASA apparently did not have time to do," Barr said. For one thing, by looking within the real-time operating system, the experts identified "unprotected critical variables." They obtained and reviewed the source code for the "sub-CPU," and they "uncovered gaps and defects in the throttle fail safes."

The experts demonstrated that "the defects we found were linked to unintended acceleration through vehicle testing," Barr said. "We also obtained and reviewed the source code for the black box and found that it can record false information about the driver's actions in the final seconds before a crash."

Stack overflow and software bugs led to memory corruption, he said. And it turns out that the crux of the issue was these memory corruptions, which acted "like ricocheting bullets."

Barr also said more than half the dozens of tasks' deaths studied by the experts in their experiments "were not detected by any fail safe."

Bookout Trial Reporting

http://www.eetimes.com/document.asp?doc_id=1319903&page_number=1 (excerpts)

"Task X death in combination with other task deaths"

14

Software is Critical: Privacy

Equifax security breach impacts 145.5 million

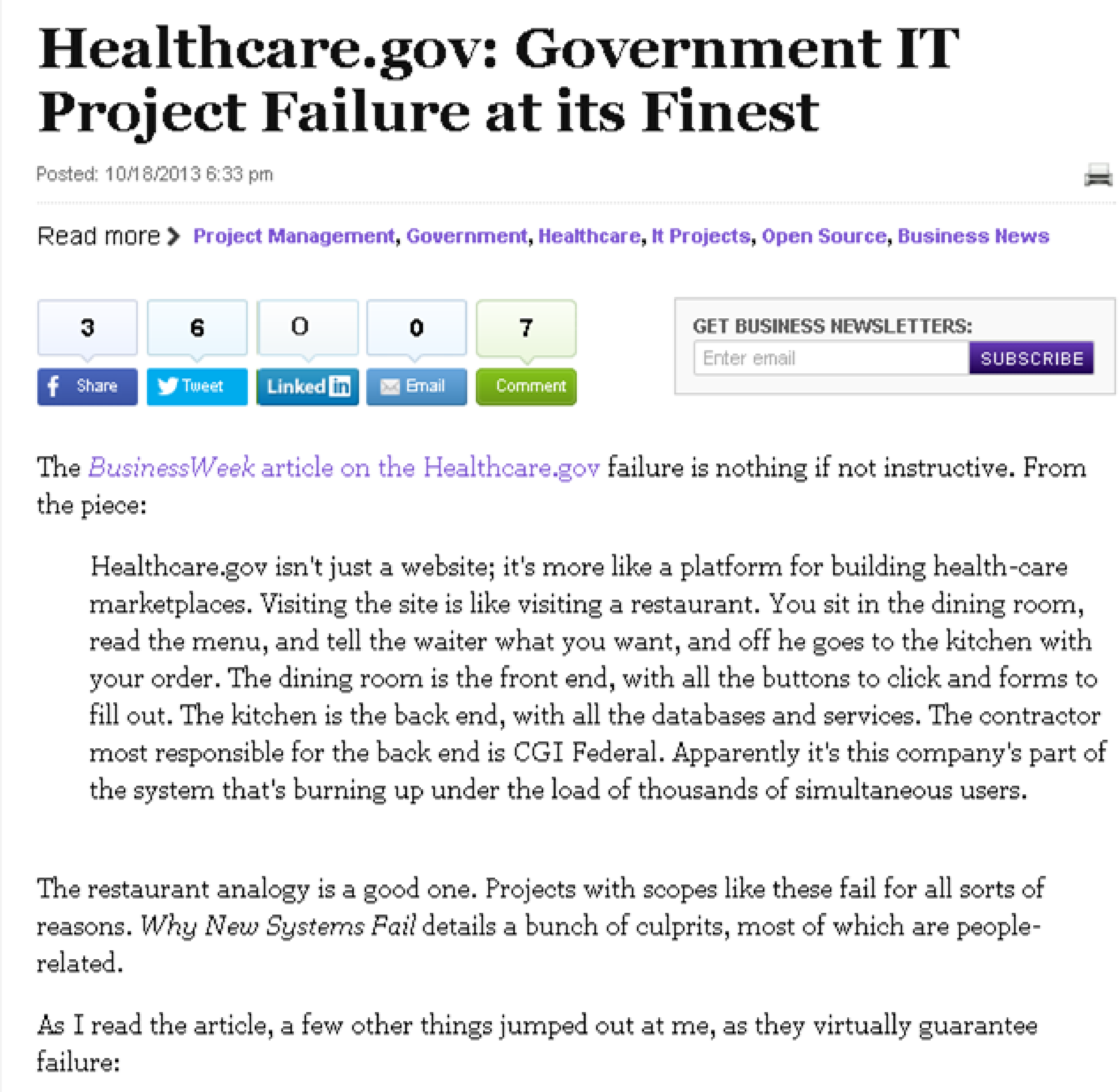
Name, SSN, DOB, Address. Also DL# and CC#

“I didn't have to do anything fancy,” the researcher told Motherboard, explaining that the site was vulnerable to a basic “forced browsing” bug. The researcher requested anonymity out of professional concerns. “**All you had to do was put in a search term** and get millions of results, just instantly—in cleartext, through a web app,” they said. In total, the researcher downloaded the data of hundreds of thousands of Americans in order to show Equifax the vulnerabilities within its systems. They said they could have downloaded the data of all of Equifax's customers in 10 minutes: “I've seen a lot of bad things, but not this bad.”

Software is Critical: Space

- The European Space Agency's Ariane 5 Flight 501 was **destroyed** 40 seconds after takeoff (June 4, 1996). The US $1 billion prototype rocket **self-destructed** due to a bug in the on-board guidance software.
- (The bug? Bad conversion of `double` to `short`, leading to an overflow.)

Software is Critical: Healthcare



Healthcare.gov: Government IT Project Failure at its Finest

Posted: 10/18/2013 6:33 pm

Read more > Project Management, Government, Healthcare, It Projects, Open Source, Business News

The *BusinessWeek* article on the Healthcare.gov failure is nothing if not instructive. From the piece:

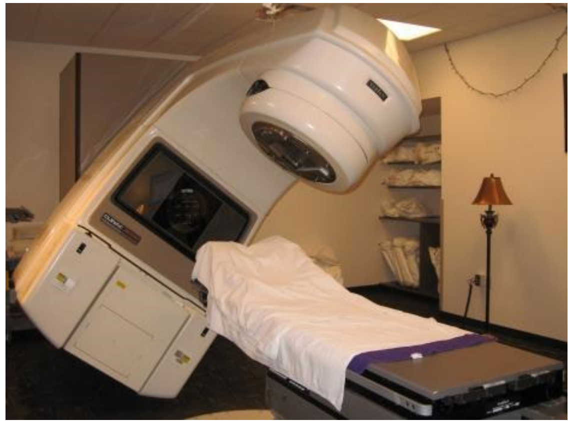
> Healthcare.gov isn't just a website; it's more like a platform for building health-care marketplaces. Visiting the site is like visiting a restaurant. You sit in the dining room, read the menu, and tell the waiter what you want, and off he goes to the kitchen with your order. The dining room is the front end, with all the buttons to click and forms to fill out. The kitchen is the back end, with all the databases and services. The contractor most responsible for the back end is CGI Federal. Apparently it's this company's part of the system that's burning up under the load of thousands of simultaneous users.

The restaurant analogy is a good one. Projects with scopes like these fail for all sorts of reasons. *Why New Systems Fail* details a bunch of culprits, most of which are people-related.

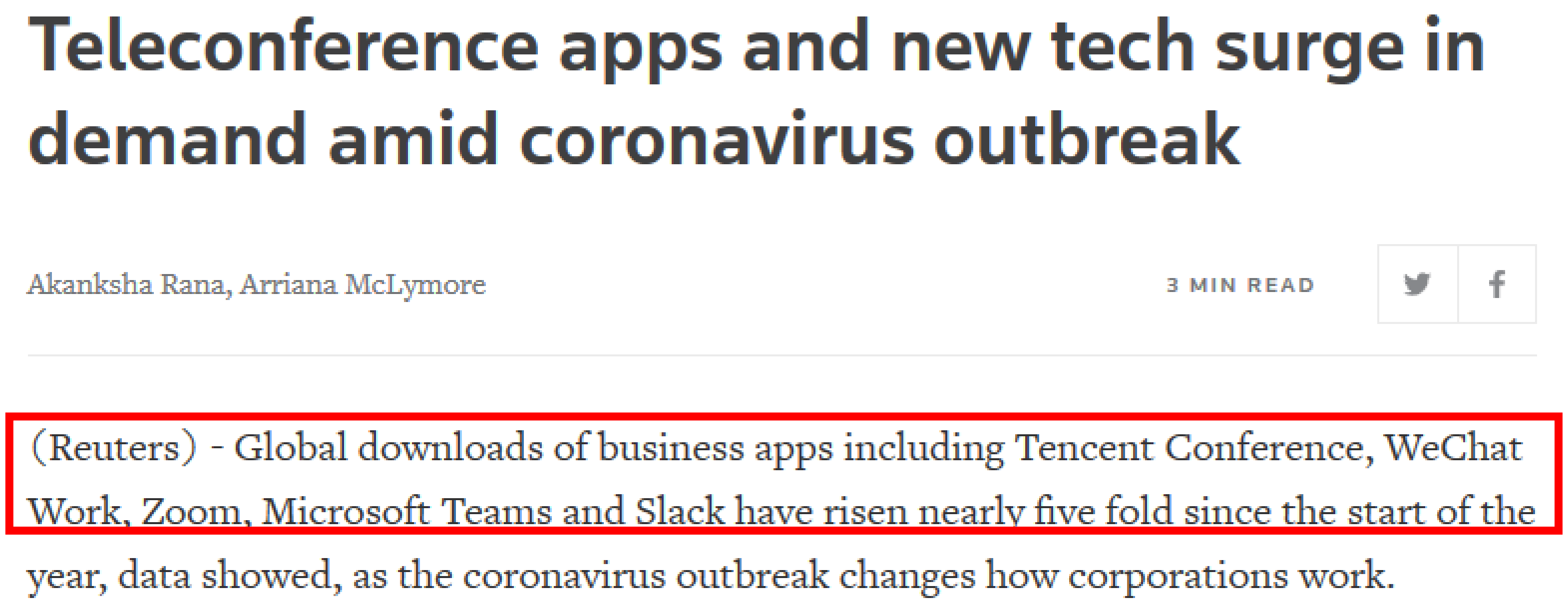
As I read the article, a few other things jumped out at me, as they virtually guarantee failure:

Software is Critical: Healthcare (!)

- Therac-25 radiation therapy machine
- At least six accidents in which patients were given massive overdoses of radiation
- Because of concurrent programming errors, it sometimes gave its patients radiation doses that were hundreds of times greater than normal, resulting in death or serious injury



Software is Critical: Coronavirus

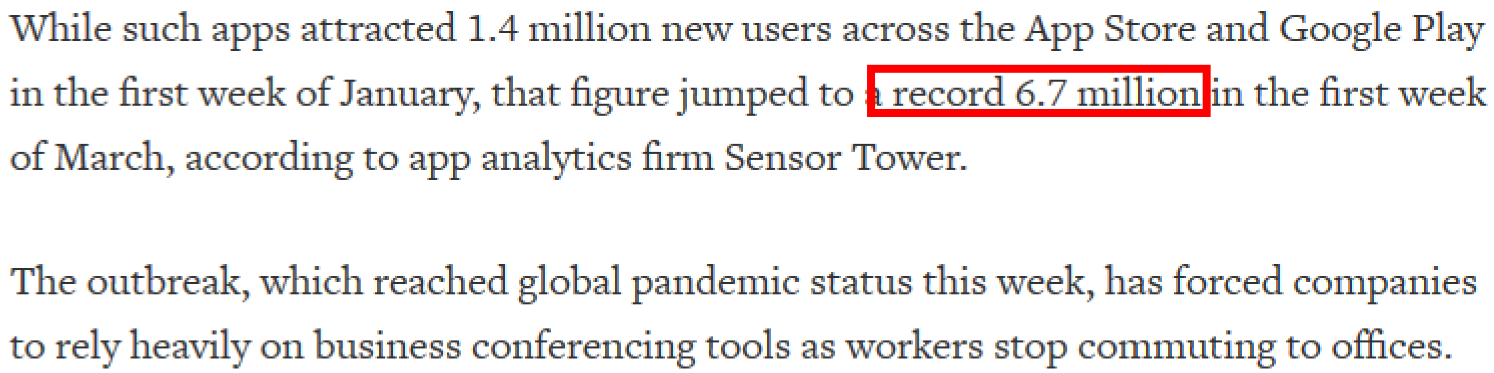


Teleconference apps and new tech surge in demand amid coronavirus outbreak

Akanksha Rana, Arriana McLymore

3 MIN READ

(Reuters) - Global downloads of business apps including Tencent Conference, WeChat Work, Zoom, Microsoft Teams and Slack have risen nearly five fold since the start of the year, data showed, as the coronavirus outbreak changes how corporations work.



While such apps attracted 1.4 million new users across the App Store and Google Play in the first week of January, that figure jumped to a record 6.7 million in the first week of March, according to app analytics firm Sensor Tower.

The outbreak, which reached global pandemic status this week, has forced companies to rely heavily on business conferencing tools as workers stop commuting to offices.

What Is Software Engineering?

What Is Software Engineering?

- “My favorite operational definition of engineering is '**design under constraint**.' Engineering is creating, designing what can be, but it is constrained by nature, by cost, by concerns of safety, reliability, environmental impact, manufacturability, maintainability, and many other such 'ilities.'”

[Bill Wulf, NAE President, The Urgency of Engineering Education Reform, 2008]

- “[Software Engineering is] The Establishment and use of sound **engineering principles** in order to obtain **economically** software that is **reliable** and works **efficiently** on **real** machines.”

[Fritz Bauer 1975, S. 524]

What is Software Engineering Research?

- “My favorite operational definition of engineering is '**design under constraint**.' Engineering is creating, designing what can be, but it is constrained by nature, by cost, by concerns of safety, reliability, environmental impact, manufacturability, maintainability, and many other such 'ilities.'”

[Bill Wulf, NAE President, The Urgency of Engineering Education Reform, 2008]

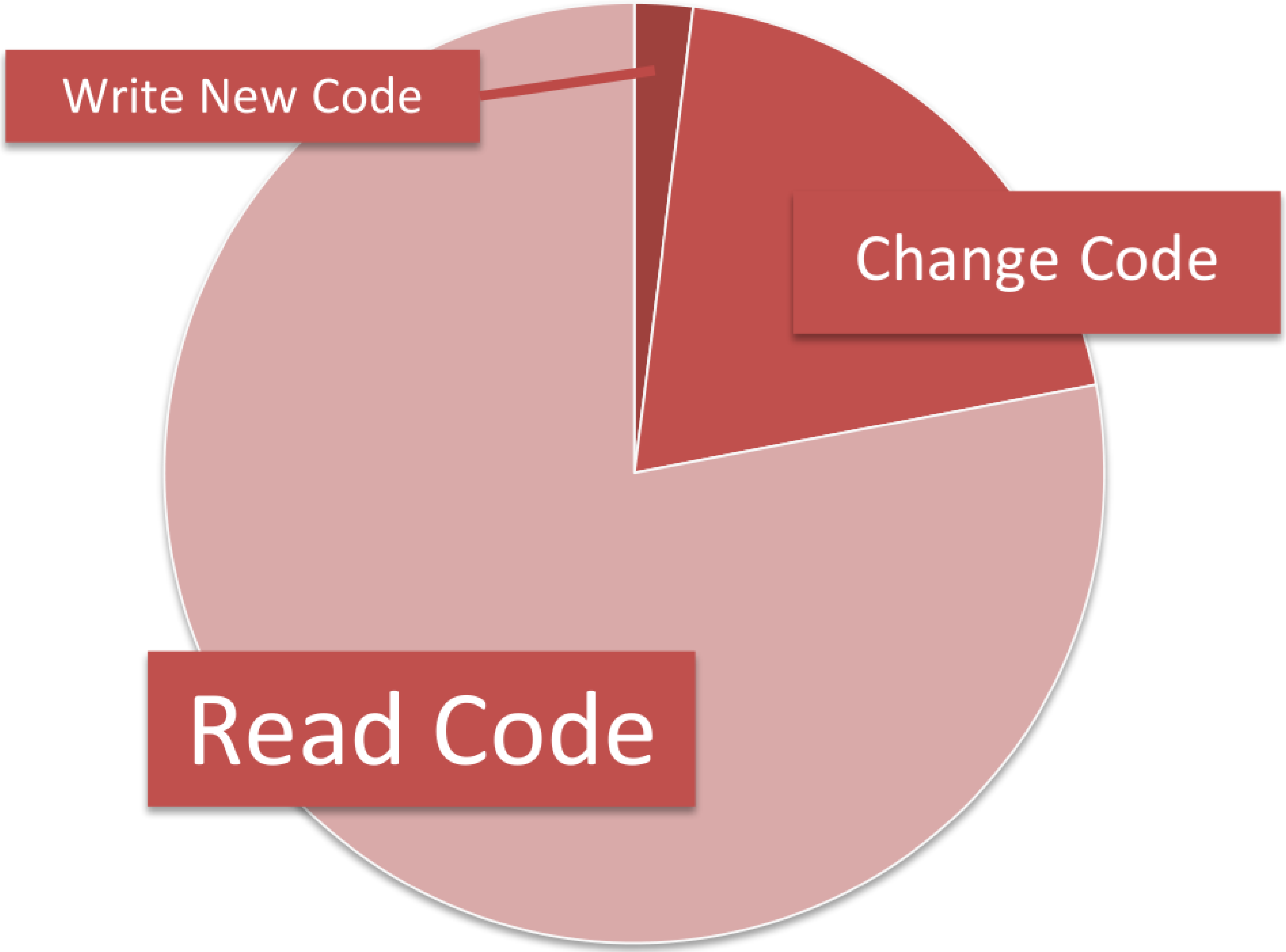
Everything that is involved in the entire process!

- “[Software Engineering is] The Establishment and use of sound **engineering principles** in order to obtain **economically** software that is **reliable** and works **efficiently** on **real** machines.”

[Fritz Bauer 1975, S. 524]

What Is Software Engineering?

- The majority of industrial software engineering is *not* writing code.
- The dominant activities in software engineering are **comprehension** and **maintenance**.



“**Understanding code** is **by far** the activity at which professional developers spend most of their time.”

[Peter Hallam. *What Do Programmers Really Do Anyway?* Microsoft.]

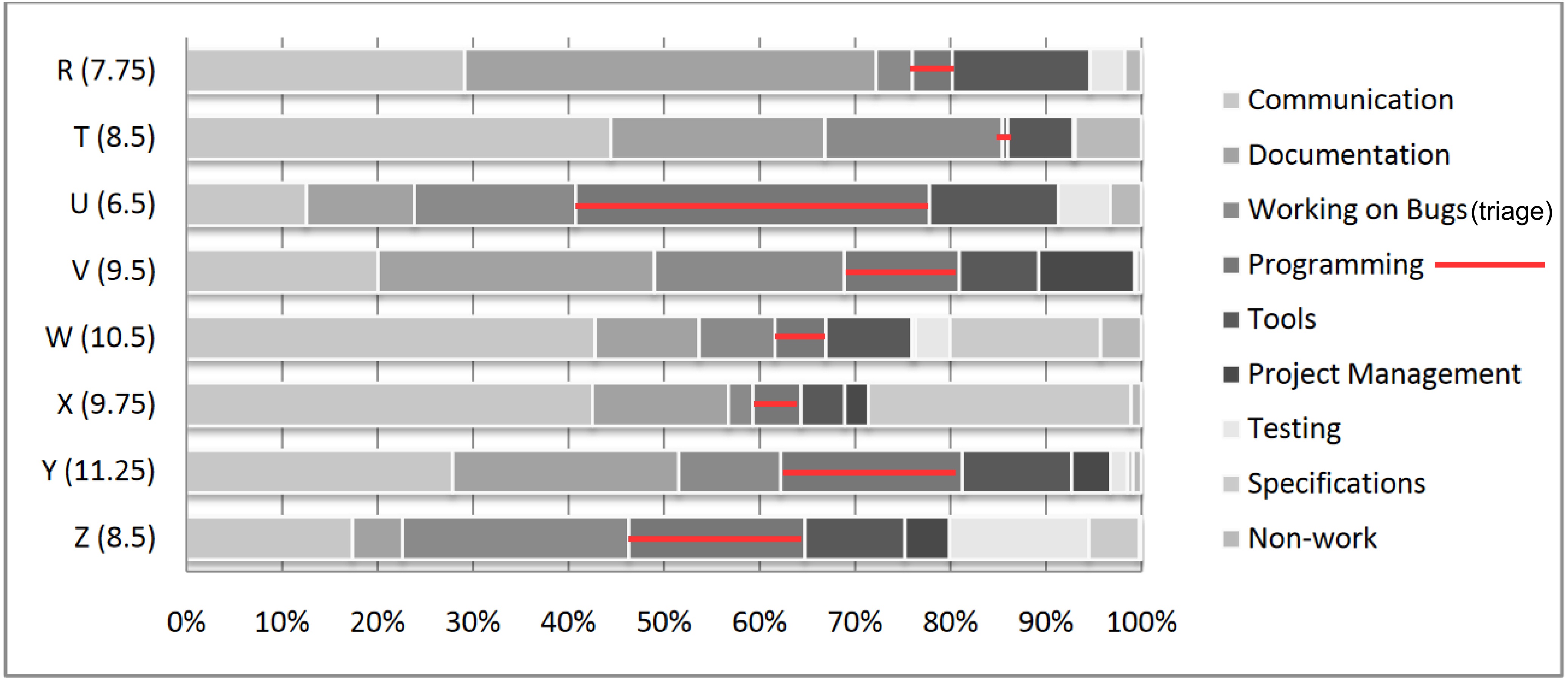
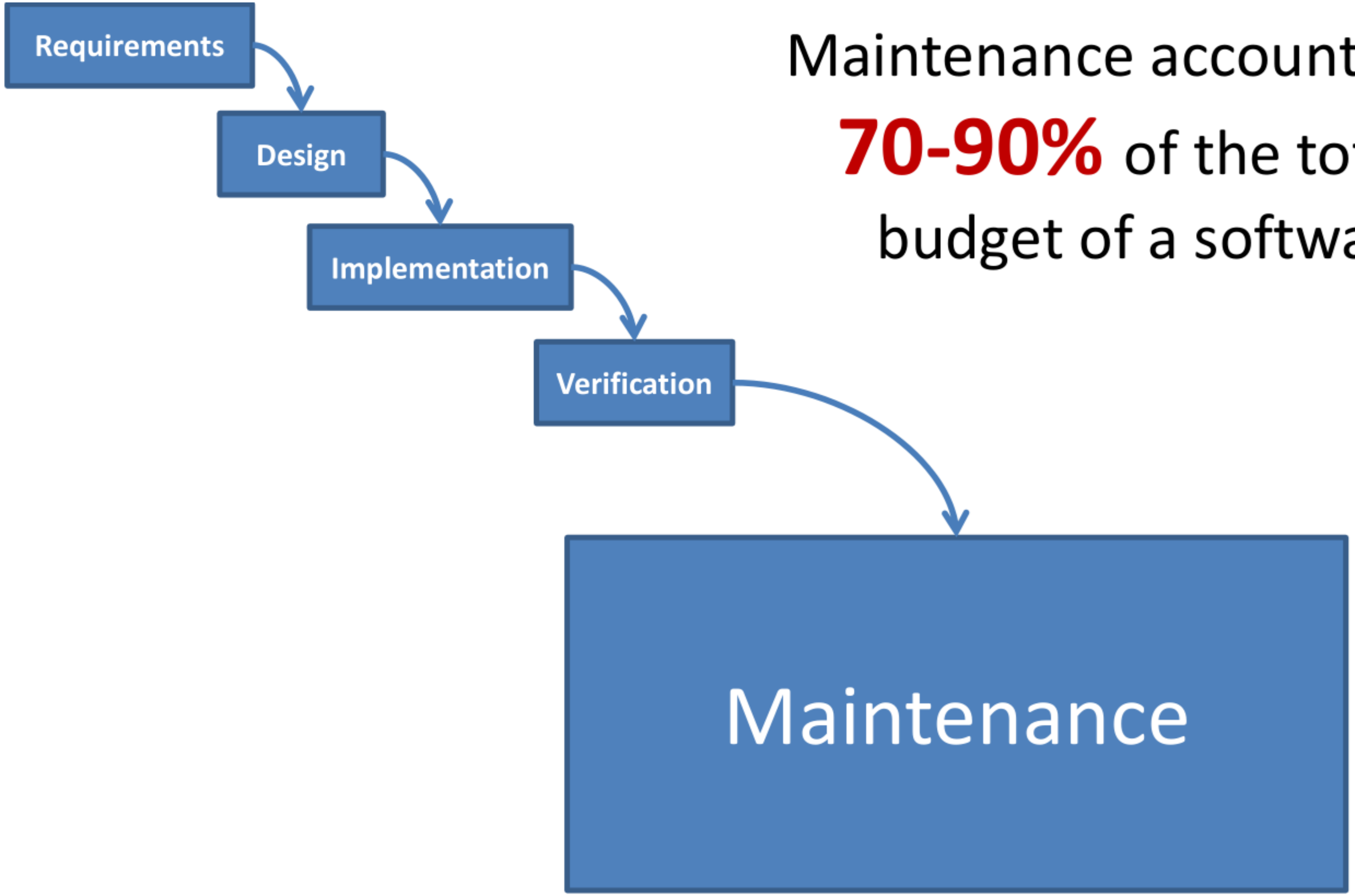


Figure 1. Tasks by time for *each* subject, normalized by the total time + time where events overlapped in each observation. Total observation time in hours is listed in parentheses after each subject's identification letter.

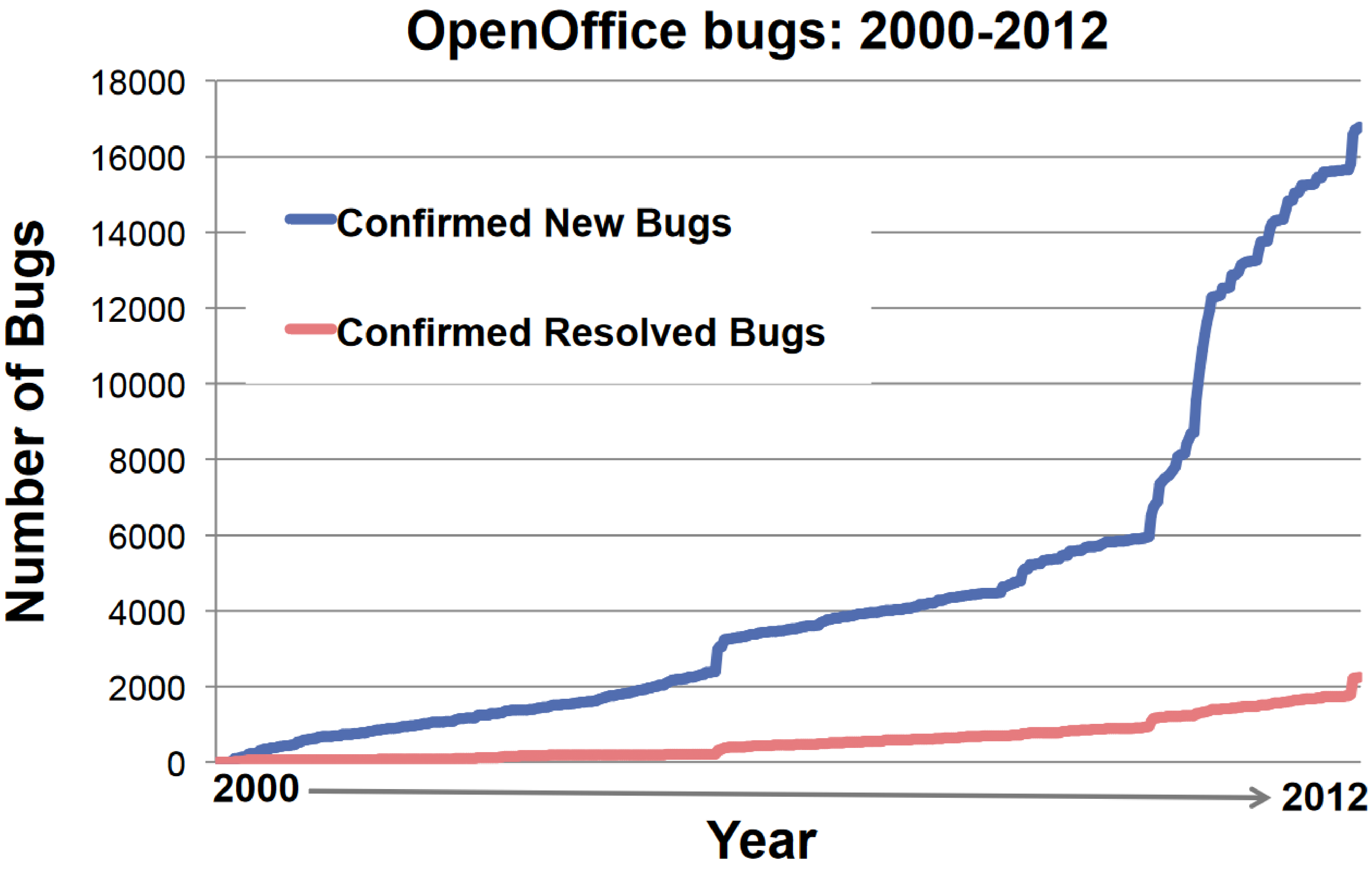
- Hour logging of new devs (1-7 months) at Microsoft: programming is 10-20% of the time.

[Begel and Simon. *Novice Software Developers, All Over Again.* Computing Education Research, September 2008. Microsoft.]



Maintenance accounts for about **70-90%** of the total lifecycle budget of a software project.

[Pigoski. *Practical Software Maintenance: Best Practices for Managing Your Software Investment*. Seacord, Plakosh, and Lewis. *Modernizing Legacy Systems: Software Technologies*.]



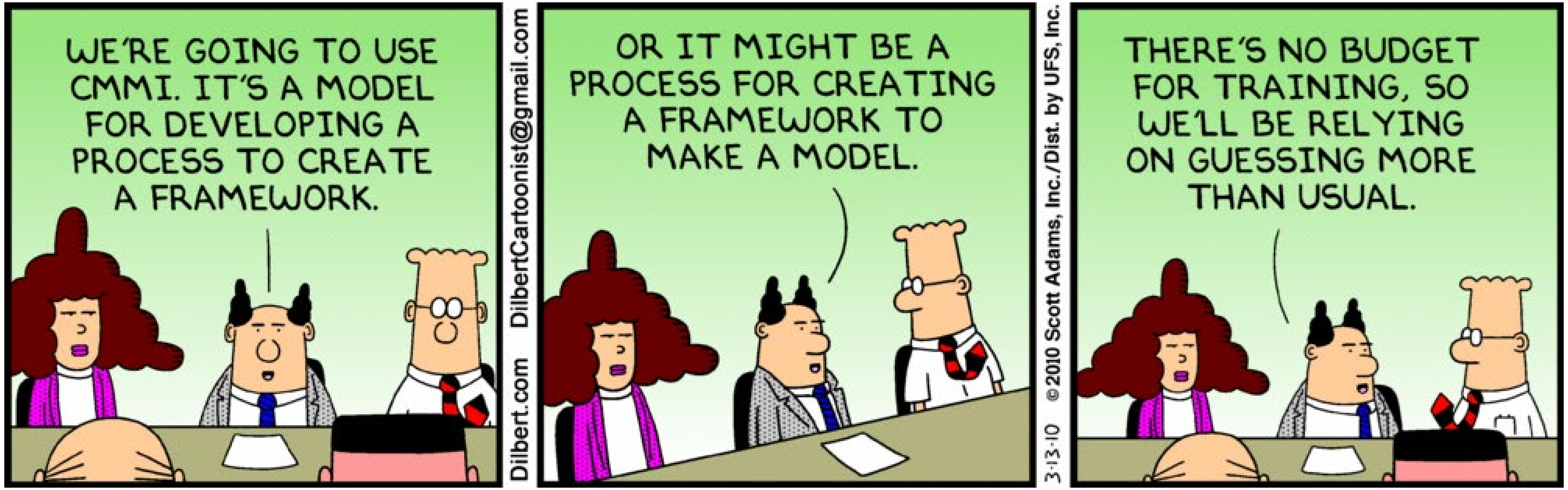
[Fry. *Leveraging Light-Weight Analyses to Aid Software Maintenance*. 2014]

A Key Issue

“Half of software engineering research is crap.”

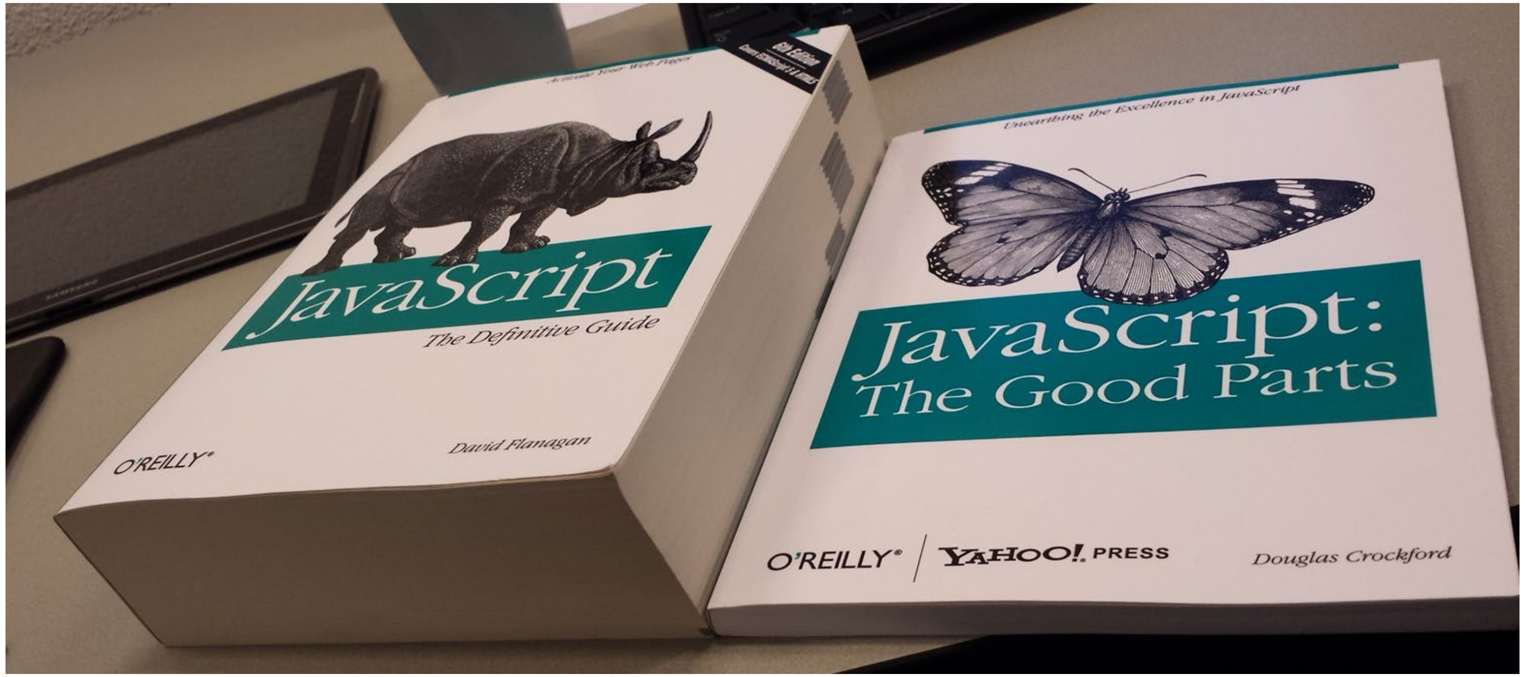
– Prof. Weimer

(Prof. Huang thinks it’s more like two-thirds crap)


WE'RE GOING TO USE CMMI. IT'S A MODEL FOR DEVELOPING A PROCESS TO CREATE A FRAMEWORK.
Dilbert.com DilbertCartoonist@gmail.com
OR IT MIGHT BE A PROCESS FOR CREATING A FRAMEWORK TO MAKE A MODEL.
3·13·10 ©2010 Scott Adams, Inc./Dist. by UFS, Inc.
THERE'S NO BUDGET FOR TRAINING, SO WE'LL BE RELYING ON GUESSING MORE THAN USUAL.

Class Philosophy

Anyway, here's the 'good parts' version.



CS8395: Upper Level Special Topic Research Course

- Introduction on SE research
- Basic research methodologies
- Practice on research activities

- Major Design Experience
- Capstone
- Industrial practice

This Seminar Course

- https://huang.isis.vanderbilt.edu/cs8395
- **Lectures + students' presentations**

Software Engineering

↓

Programs + Programmers (+ Clients + Users)

Core Course Topics

- Program analysis
  - Is a program correct? Does a program have certain property?
- Program testing and repair
  - How to effectively conduct this boring but extremely important task?
- Software security
  - Vulnerability in programs
- Human factors in SE
  - SE is a human activity, what can go wrong?
- AI for SE; SE for AI
  - Leverage AI tools to automate SE tasks
  - Optimizing AI as a special SE process
- Open source software
  - Global collaboration, diverse applications, unique challenges

Research themes

- How to combine SE with X?
  - X can be your own research direction
    - PL
    - HCI
    - Hardware
    - Security
    - Cyber Physical System
    - Social Science
    - Psychology
    - Neuroscience
    - Education
    - …

More on next lecture

Computer Science

- *This* English mathematician and writer published the first algorithm (~1842) to be carried out by a general-purpose computer and is often called the first computer programmer.

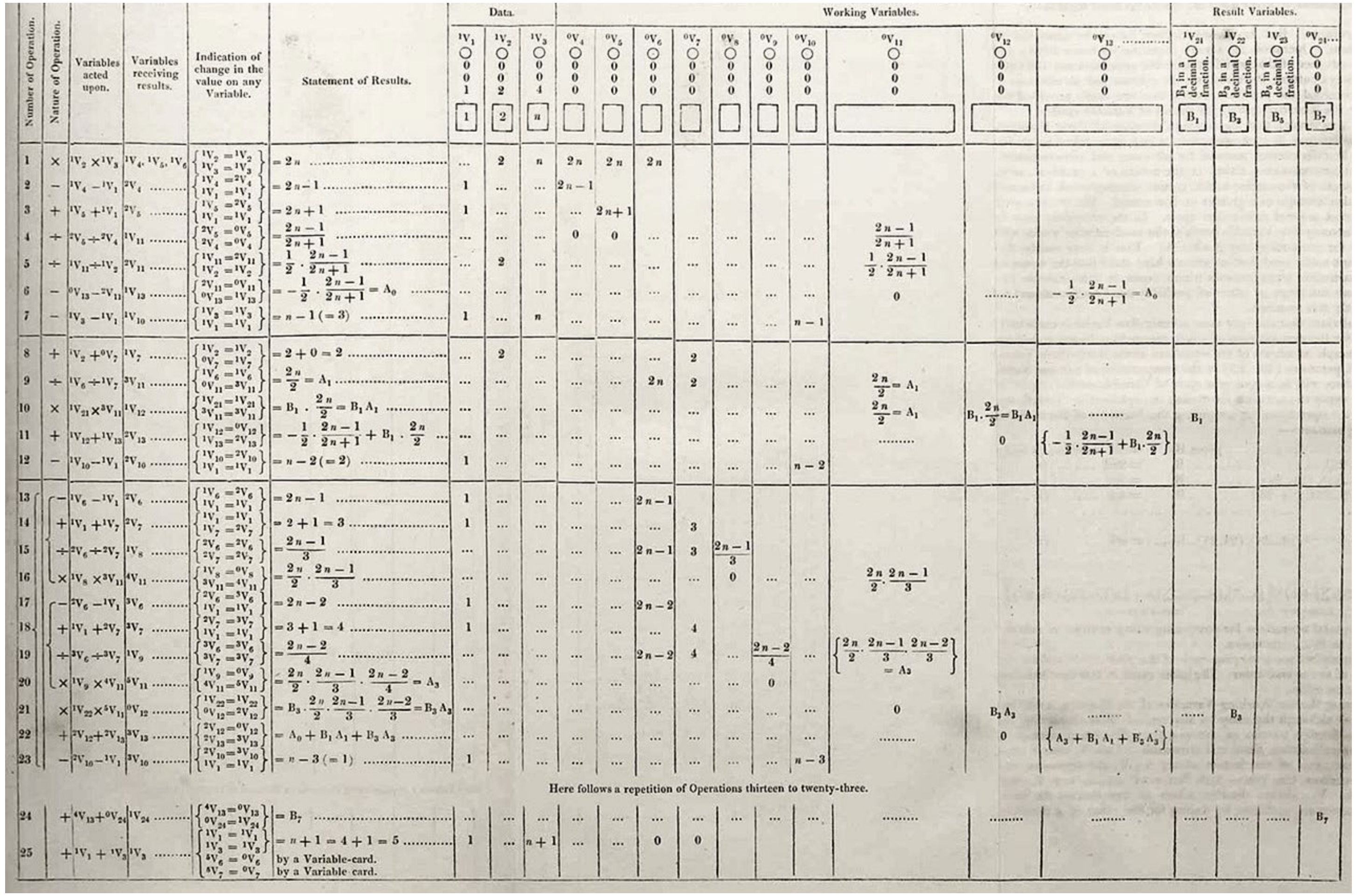
Computer Science

- *This* English mathematician and writer published the first algorithm (~1842) to be carried out by a general-purpose computer and is often called the first computer programmer.

Ada Lovelace

Computer Science

- *What* did that first program do?



Number of Operation.	Nature of Operation.	Variables acted upon.	Variables receiving results.	Indication of change in the value on any Variable.	Statement of Results.	Data.			Working Variables.										Result Variables.			
						$^1V_1$ ○ 0 0 0 1 1	$^1V_2$ ○ 0 0 0 2 2	$^1V_3$ ○ 0 0 0 4 n	$^0V_4$ ○ 0 0 0 0	$^0V_5$ ○ 0 0 0 0	$^0V_6$ ○ 0 0 0 0	$^0V_7$ ○ 0 0 0 0	$^0V_8$ ○ 0 0 0 0	$^0V_9$ ○ 0 0 0 0	$^0V_{10}$ ○ 0 0 0 0	$^0V_{11}$ ○ 0 0 0 0	$^0V_{12}$ ○ 0 0 0 0	$^0V_{13}$ .... ○ 0 0 0 0	$^1V_{21}$ ○ B_1 in a decimal fraction. B_1	$^1V_{22}$ ○ B_3 in a decimal fraction. B_3	$^1V_{23}$ ○ B_5 in a decimal fraction. B_5	$^0V_{24}$... ○ 0 0 0 B_7
1	×	$^1V_2 \times {}^1V_3$	$^1V_4, {}^1V_5, {}^1V_6$	$\begin{Bmatrix} {}^1V_2 = {}^1V_2 \\ {}^1V_3 = {}^1V_3 \end{Bmatrix}$	$= 2n$	...	2	n	$2n$	$2n$	$2n$											
2	−	$^1V_4 - {}^1V_1$	$^2V_4$	$\begin{Bmatrix} {}^1V_4 = {}^2V_4 \\ {}^1V_1 = {}^1V_1 \end{Bmatrix}$	$= 2n - 1$	1	...	...	$2n - 1$													
3	+	$^1V_5 + {}^1V_1$	$^2V_5$	$\begin{Bmatrix} {}^1V_5 = {}^2V_5 \\ {}^1V_1 = {}^1V_1 \end{Bmatrix}$	$= 2n + 1$	1	...	...	...	$2n + 1$												
4	÷	$^2V_5 \div {}^2V_4$	$^1V_{11}$	$\begin{Bmatrix} {}^2V_5 = {}^0V_5 \\ {}^2V_4 = {}^0V_4 \end{Bmatrix}$	$= \frac{2n-1}{2n+1}$	...	...	...	0	0	...	...	...	...	...	$\frac{2n-1}{2n+1}$						
5	÷	$^1V_{11} \div {}^1V_2$	$^2V_{11}$	$\begin{Bmatrix} {}^1V_{11} = {}^2V_{11} \\ {}^1V_2 = {}^1V_2 \end{Bmatrix}$	$= \frac{1}{2} \cdot \frac{2n-1}{2n+1}$	...	2	...	...	...	...	...	...	...	...	$\frac{1}{2} \cdot \frac{2n-1}{2n+1}$						
6	−	$^0V_{13} - {}^2V_{11}$	$^1V_{13}$	$\begin{Bmatrix} {}^2V_{11} = {}^0V_{11} \\ {}^0V_{13} = {}^1V_{13} \end{Bmatrix}$	$= -\frac{1}{2} \cdot \frac{2n-1}{2n+1} = A_0$	...	...	...	...	...	...	...	...	...	...	0	....	$-\frac{1}{2} \cdot \frac{2n-1}{2n+1} = A_0$				
7	−	$^1V_3 - {}^1V_1$	$^1V_{10}$	$\begin{Bmatrix} {}^1V_3 = {}^1V_3 \\ {}^1V_1 = {}^1V_1 \end{Bmatrix}$	$= n - 1 \ (= 3)$	1	...	n	...	...	...	...	...	...	$n - 1$							
8	+	$^1V_2 + {}^0V_7$	$^1V_7$	$\begin{Bmatrix} {}^1V_2 = {}^1V_2 \\ {}^0V_7 = {}^1V_7 \end{Bmatrix}$	$= 2 + 0 = 2$	...	2	...	...	...	...	2										
9	÷	$^1V_6 \div {}^1V_7$	$^3V_{11}$	$\begin{Bmatrix} {}^1V_6 = {}^1V_6 \\ {}^0V_{11} = {}^3V_{11} \end{Bmatrix}$	$= \frac{2n}{2} = A_1$	...	...	...	...	...	$2n$	2	...	...	...	$\frac{2n}{2} = A_1$						
10	×	$^1V_{21} \times {}^3V_{11}$	$^1V_{12}$	$\begin{Bmatrix} {}^1V_{21} = {}^1V_{21} \\ {}^3V_{11} = {}^3V_{11} \end{Bmatrix}$	$= B_1 \cdot \frac{2n}{2} = B_1 A_1$	...	...	...	...	...	...	...	...	...	...	$\frac{2n}{2} = A_1$	$B_1 \cdot \frac{2n}{2} = B_1 A_1$	....	B_1			
11	+	$^1V_{12} + {}^1V_{13}$	$^2V_{13}$	$\begin{Bmatrix} {}^1V_{12} = {}^0V_{12} \\ {}^1V_{13} = {}^2V_{13} \end{Bmatrix}$	$= -\frac{1}{2} \cdot \frac{2n-1}{2n+1} + B_1 \cdot \frac{2n}{2}$ ...	...	...	...	...	...	...	...	...	...	...	....	0	$\left\{-\frac{1}{2} \cdot \frac{2n-1}{2n+1} + B_1 \cdot \frac{2n}{2}\right\}$				
12	−	$^1V_{10} - {}^1V_1$	$^2V_{10}$	$\begin{Bmatrix} {}^1V_{10} = {}^2V_{10} \\ {}^1V_1 = {}^1V_1 \end{Bmatrix}$	$= n - 2 \ (= 2)$	1	...	...	...	...	...	...	...	...	$n - 2$							
13	−	$^1V_6 - {}^1V_1$	$^2V_6$	$\begin{Bmatrix} {}^1V_6 = {}^2V_6 \\ {}^1V_1 = {}^1V_1 \end{Bmatrix}$	$= 2n - 1$	1	...	...	...	...	$2n - 1$											
14	+	$^1V_1 + {}^1V_7$	$^2V_7$	$\begin{Bmatrix} {}^1V_1 = {}^1V_1 \\ {}^1V_7 = {}^2V_7 \end{Bmatrix}$	$= 2 + 1 = 3$	1	...	...	...	...	...	3										
15	÷	$^2V_6 \div {}^2V_7$	$^1V_8$	$\begin{Bmatrix} {}^2V_6 = {}^2V_6 \\ {}^2V_7 = {}^2V_7 \end{Bmatrix}$	$= \frac{2n-1}{3}$	...	...	...	...	...	$2n - 1$	3	$\frac{2n-1}{3}$									
16	×	$^1V_8 \times {}^3V_{11}$	$^4V_{11}$	$\begin{Bmatrix} {}^1V_8 = {}^0V_8 \\ {}^3V_{11} = {}^4V_{11} \end{Bmatrix}$	$= \frac{2n}{2} \cdot \frac{2n-1}{3}$	...	...	...	...	...	...	...	0	...	...	$\frac{2n}{2} \cdot \frac{2n-1}{3}$						
17	−	$^2V_6 - {}^1V_1$	$^3V_6$	$\begin{Bmatrix} {}^2V_6 = {}^3V_6 \\ {}^1V_1 = {}^1V_1 \end{Bmatrix}$	$= 2n - 2$	1	...	...	...	...	$2n - 2$											
18	+	$^1V_1 + {}^2V_7$	$^3V_7$	$\begin{Bmatrix} {}^2V_7 = {}^3V_7 \\ {}^1V_1 = {}^1V_1 \end{Bmatrix}$	$= 3 + 1 = 4$	1	...	...	...	...	...	4										
19	÷	$^3V_6 \div {}^3V_7$	$^1V_9$	$\begin{Bmatrix} {}^3V_6 = {}^3V_6 \\ {}^3V_7 = {}^3V_7 \end{Bmatrix}$	$= \frac{2n-2}{4}$	...	...	...	...	...	$2n - 2$	4	...	$\frac{2n-2}{4}$	...	$\left\{\frac{2n}{2} \cdot \frac{2n-1}{3} \cdot \frac{2n-2}{3} = A_3\right\}$						
20	×	$^1V_9 \times {}^4V_{11}$	$^5V_{11}$	$\begin{Bmatrix} {}^1V_9 = {}^0V_9 \\ {}^4V_{11} = {}^5V_{11} \end{Bmatrix}$	$= \frac{2n}{2} \cdot \frac{2n-1}{3} \cdot \frac{2n-2}{4} = A_3$ ...	...	...	...	...	...	...	...	...	0								
21	×	$^1V_{22} \times {}^5V_{11}$	$^0V_{12}$	$\begin{Bmatrix} {}^1V_{22} = {}^1V_{22} \\ {}^0V_{12} = {}^2V_{12} \end{Bmatrix}$	$= B_3 \cdot \frac{2n}{2} \cdot \frac{2n-1}{3} \cdot \frac{2n-2}{3} = B_3 A_3$	...	...	...	...	...	...	...	...	...	...	0	$B_3 A_3$	....	......	B_3		
22	+	$^2V_{12} + {}^2V_{13}$	$^3V_{13}$	$\begin{Bmatrix} {}^2V_{12} = {}^0V_{12} \\ {}^2V_{13} = {}^3V_{13} \end{Bmatrix}$	$= A_0 + B_1 A_1 + B_3 A_3$	...	...	...	...	...	...	...	...	...	...	....	0	$\{A_0 + B_1 A_1 + B_3 A_3\}$				
23	−	$^2V_{10} - {}^1V_1$	$^3V_{10}$	$\begin{Bmatrix} {}^2V_{10} = {}^3V_{10} \\ {}^1V_1 = {}^1V_1 \end{Bmatrix}$	$= n - 3 \ (= 1)$	1	...	...	...	...	...	...	...	...	$n - 3$							
					Here follows a repetition of Operations thirteen to twenty-three.																	
24	+	$^4V_{13} + {}^0V_{24}$	$^1V_{24}$	$\begin{Bmatrix} {}^4V_{13} = {}^0V_{13} \\ {}^0V_{24} = {}^1V_{24} \end{Bmatrix}$	$= B_7$	...	...	...	...	...	...	...	...	...	...	....	....	....	......	......	......	B_7
25	+	$^1V_1 + {}^1V_3$	$^1V_3$	$\begin{Bmatrix} {}^1V_1 = {}^1V_1 \\ {}^1V_3 = {}^1V_3 \\ {}^5V_6 = {}^0V_6 \\ {}^5V_7 = {}^0V_7 \end{Bmatrix}$	$= n + 1 = 4 + 1 = 5$ .... by a Variable-card. by a Variable card.	1	...	$n + 1$	...	...	0	0										

Who should take this class?

- If you want to learn:
    - What do we care about in SE research?
    - What methodologies are used in SE research?
    - How to evaluate and design a research project in SE?
    - How can you apply SE methods to your own research/work?

Who should take this class?

- If you want to learn:
  - What do we care about in SE research?
  - What methodologies are used in SE research?
  - How to evaluate and design a research project in SE?
  - How can you apply SE methods to your own research?

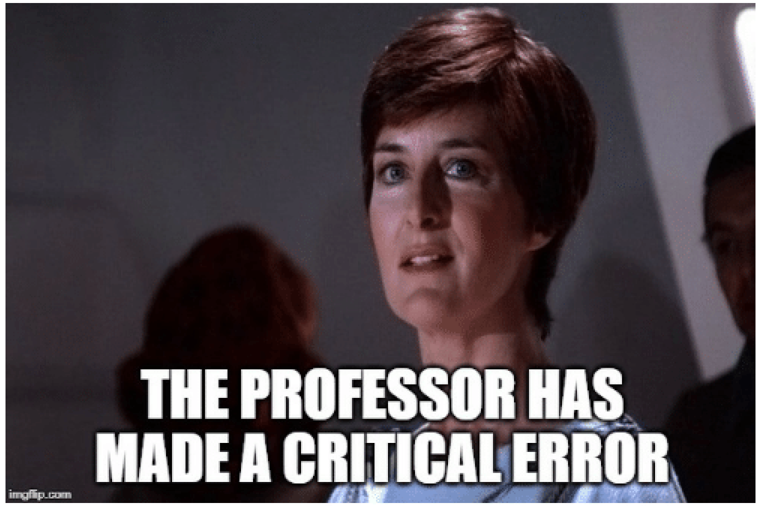
- If you want to learn skills in SE industry (e.g., daily work as software developers, coding):
  - Consider **CS4278/5278 Principles of Software Engineering** instead

How Hard Is This Course?

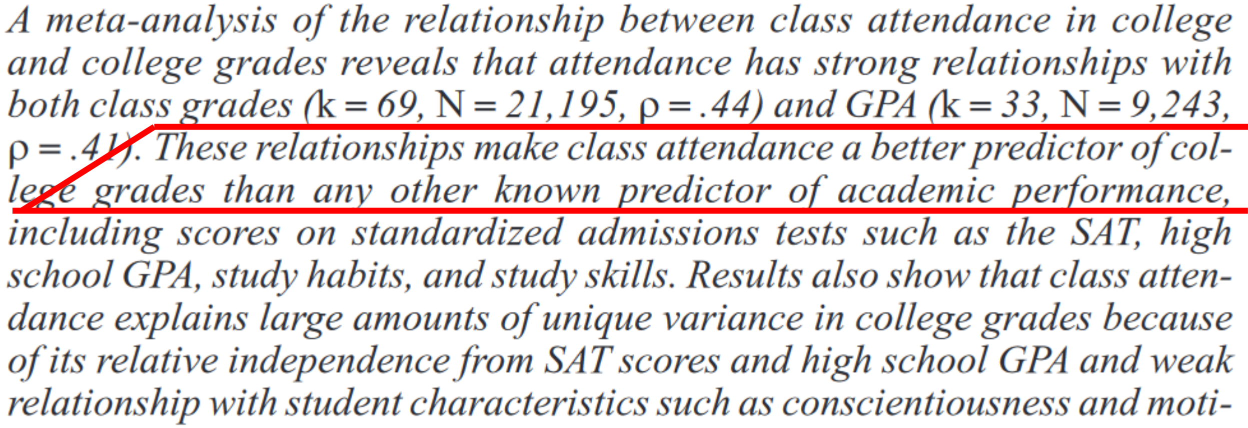
- Light
  - “Coding”
- Hard
  - Quickly learn about a new research area
  - Understand research approaches on a high level
    - Every topic included in this course can be extended to a semester-long course
  - Creative and critical thinking on a new research idea
  - How to design a good research project
  - How to formally present a research project: presentations and proposals

Attendance

- I would prefer you attend the lectures live as they are delivered
  - In-class discussion
  - Provide advice for others' work – peer review
- Students are required to evaluate each other’s presentations



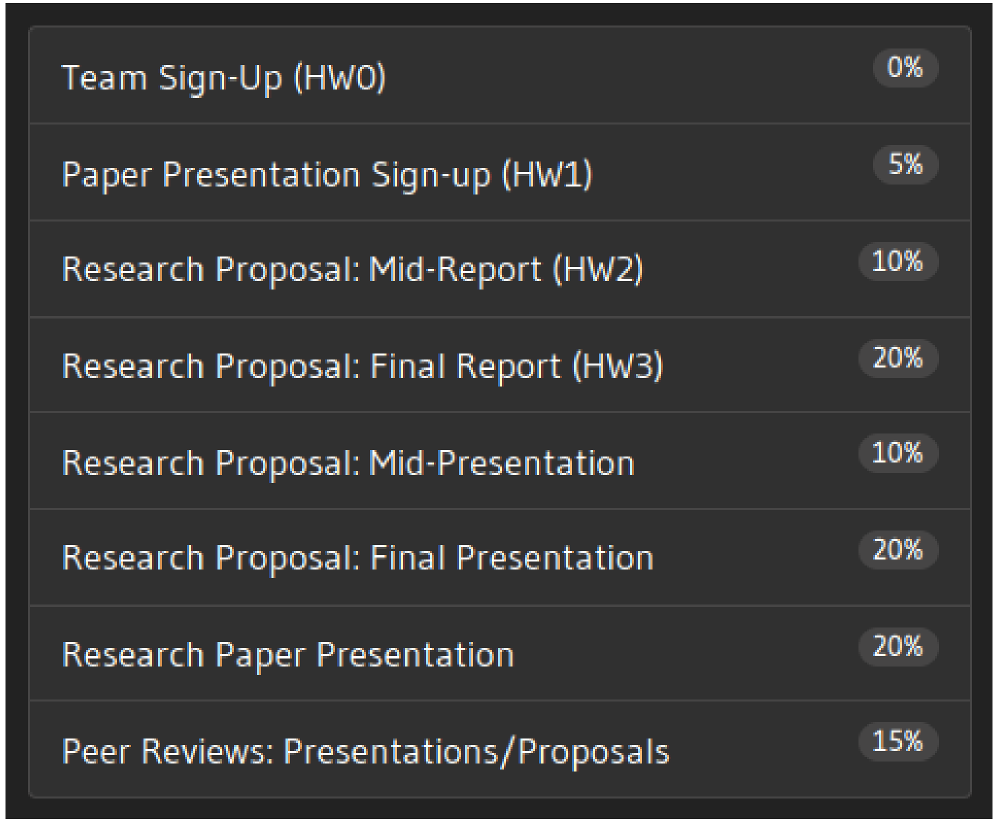
When the syllabus says they don't take attendance and participation doesn't matter



A meta-analysis of the relationship between class attendance in college and college grades reveals that attendance has strong relationships with both class grades ($k = 69$*,* $N = 21{,}195$*,* $\rho = .44$*) and GPA (*$k = 33$*,* $N = 9{,}243$*,* $\rho = .41$*). These relationships make class attendance a better predictor of college grades than any other known predictor of academic performance, including scores on standardized admissions tests such as the SAT, high school GPA, study habits, and study skills. Results also show that class attendance explains large amounts of unique variance in college grades because of its relative independence from SAT scores and high school GPA and weak relationship with student characteristics such as conscientiousness and moti-*

Crede et al. “Class Attendance in College: A Meta-Analytic Review of the Relationship of Class Attendance With Grades and Student Characteristics.” Review of Educational Research, 2010. Vol. 80. DOI: 10.3102/0034654310362998

Assignments and Grading



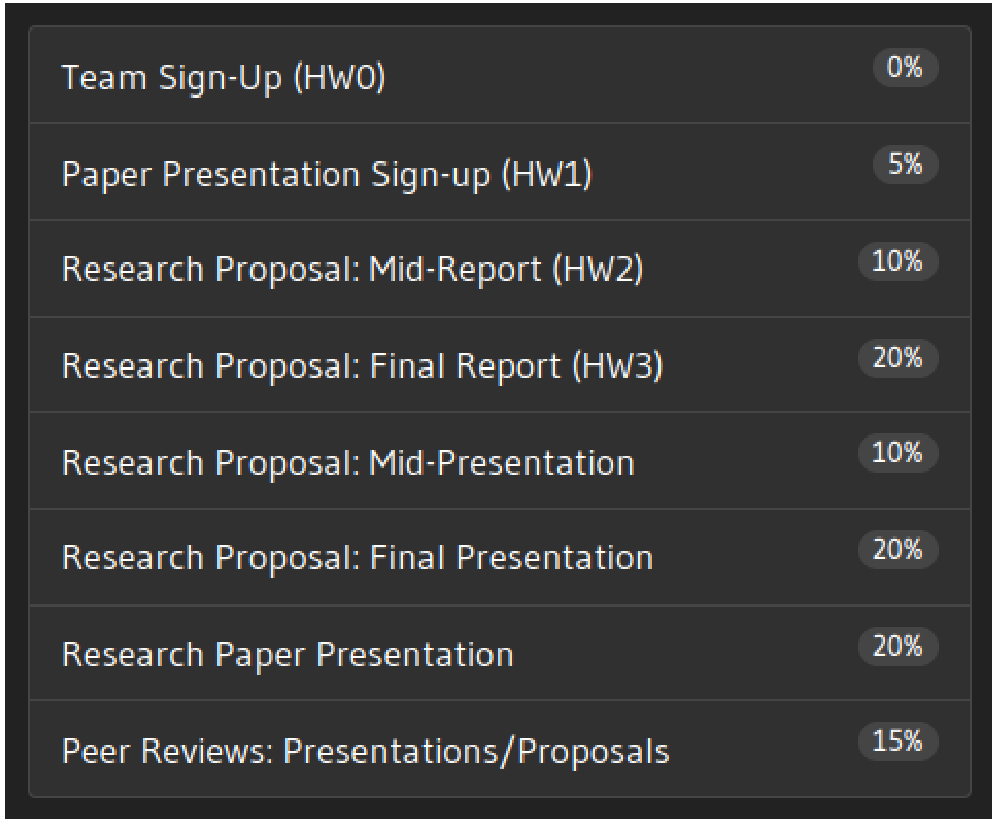
Team Sign-Up (HW0)	0%
Paper Presentation Sign-up (HW1)	5%
Research Proposal: Mid-Report (HW2)	10%
Research Proposal: Final Report (HW3)	20%
Research Proposal: Mid-Presentation	10%
Research Proposal: Final Presentation	20%
Research Paper Presentation	20%
Peer Reviews: Presentations/Proposals	15%

- HW0: does not contribute to final grades
- HW1: sign up for research paper presentation
- HW2: mid research proposal
- HW3: final research proposal
- HW report submission (HW2 and 3): Brightspace

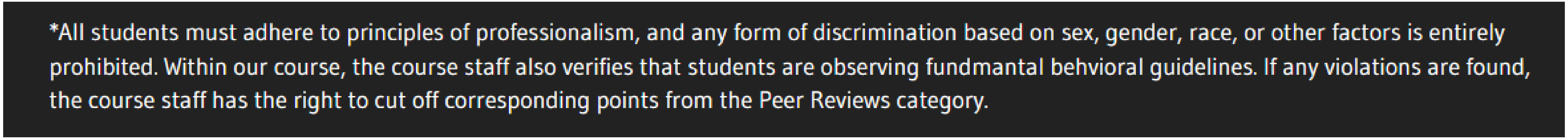
Where can you find all the information of every homework assignment?

Walk through the course website

Assignments and Grading



Team Sign-Up (HW0)	0%
Paper Presentation Sign-up (HW1)	5%
Research Proposal: Mid-Report (HW2)	10%
Research Proposal: Final Report (HW3)	20%
Research Proposal: Mid-Presentation	10%
Research Proposal: Final Presentation	20%
Research Paper Presentation	20%
Peer Reviews: Presentations/Proposals	15%



*All students must adhere to principles of professionalism, and any form of discrimination based on sex, gender, race, or other factors is entirely prohibited. Within our course, the course staff also verifies that students are observing fundmantal behvioral guidelines. If any violations are found, the course staff has the right to cut off corresponding points from the Peer Reviews category.

Readings

- No expensive, outdated textbook
- Readings for lectures
  - Regular lecture from the instructor and guest lectures
  - Paper presentations

HW0: Optional Teams

- *Modern industrial software engineering is almost exclusively team-based*
- But this is a special topic research course, not a Capstone/MDE
- You may work **alone or in teams (size of 2 or 3)** of your own choosing
  - **We are not responsible if your partner disappears**

HW0 is due on Sunday, Sep 8
(Enrollment deadline is Aug 30)

HW1: Sign up for research paper presentation

- Only if we can “finalize” the teams from HW0…
  - So I know how to best arrange the presentations: how many papers, how long, how many per course meeting, etc.
- Everything on the website will be updated and finalized after HW0 (hopefully)

Assignments: HW2 and HW3

- A research proposal
- HW2: mid-report
  - Forcing function to start early ☺
  - Discussion in class to collect advice from everyone
- HW3: full proposal
- Use the NSF proposal style
- We have a lecture to guide you on project selection

Presentations and Peer Reviews

- Presentations
  - Research paper from a list provided by the instructor
  - Your own research proposal
    - Mid
    - Final
- Peer Reviews
  - Technically this is for papers
  - But we use it for presentations in this course
  - Single blind (the instructor serves as the "chair")
  - An online form before every presentation
  - Your reviews will be anonymously sent to the presenter after the class

Assignments and Grading

- Grace period
  - One per team for written assignments only (i.e., HW2 or HW3), one extra week without penalty
  - You MUST email the course staff (both the instructor and the GSI) before the original deadline
- Late policy
  - 20% off per day, zero points after 4 days
  - Absolutely NO late policy for presentations – you get zero if you miss a presentation

Guest Lecture

- Software security
  - Prof. Kevin Leach: Malware analysis
  - I will be fully occupied with NSF panels

Schedule

- Plan 1
  - Lectures -> paper presentations -> Lectures -> paper presentations…
- Plan 2
  - All lectures -> all presentations

Research Support

- Happy to talk about research opportunities: email
- All of students' data in this class may be used in future research (**de-identified**). You will need to notify the instructor by emails if you don't want your data to be included.
  - For example, home assignments, project reports.
- Support research on human factors
  - Piazza advertisement
  - In-class advertisement
  - Will **NOT** affect your grades

Questions?

- You are responsible for all assignments at their listed times even if you are added to the course late.
  - Ask the instructor to add you to the Piazza and Brightspace if you aren’t in yet