CS 2001: Research Topics

Introduction

Prof. Adriana Kovashka
University of Pittsburgh
August 26, 2025

About the Instructor



Born 1985 in
Sofia, Bulgaria



Got BA in 2008 at
Pomona College, CA
(Computer Science &
Media Studies)



Got PhD in 2014
at University of
Texas at Austin
(Computer Vision)

Course Info

- **Course website:** http://people.cs.pitt.edu/~kovashka/cs2001_fa2025
- **Instructor**: Adriana Kovashka (kovashka@cs.pitt.edu)

  → Use "CS2001" at the beginning of your Subject
- **Office:** Sennott Square 5325
- **Class:** Tue/Thu, 11am-12:15pm
- **Office hours**: Tue/Thu, 12:15-1pm (preferred) and by appointment

Topics

1. PhD journey and logistics
2. Learning from and evaluating the literature
3. Developing and pitching ideas
4. Writing and presenting your work
5. In-class work (writing, feedback, revising)
6. Course project presentation
7. Reading group / paper presentations
8. Faculty talks

Grading

- Participation in discussions, feedback to peers, questions asked: 20%
- Paper section drafts, offering feedback, revisions, writing and responding to review: 20%
- Critical paper presentation: 10%
- Literature review presentation: 10%
- Annotated bibliography: 10%
- Idea pitch: 10%
- Paper pair presentation: 10%
- Project presentation: 10%

Policies, schedule, readings

- Course website:

Please bring your laptops!
(and only use them for class work)

Introductions

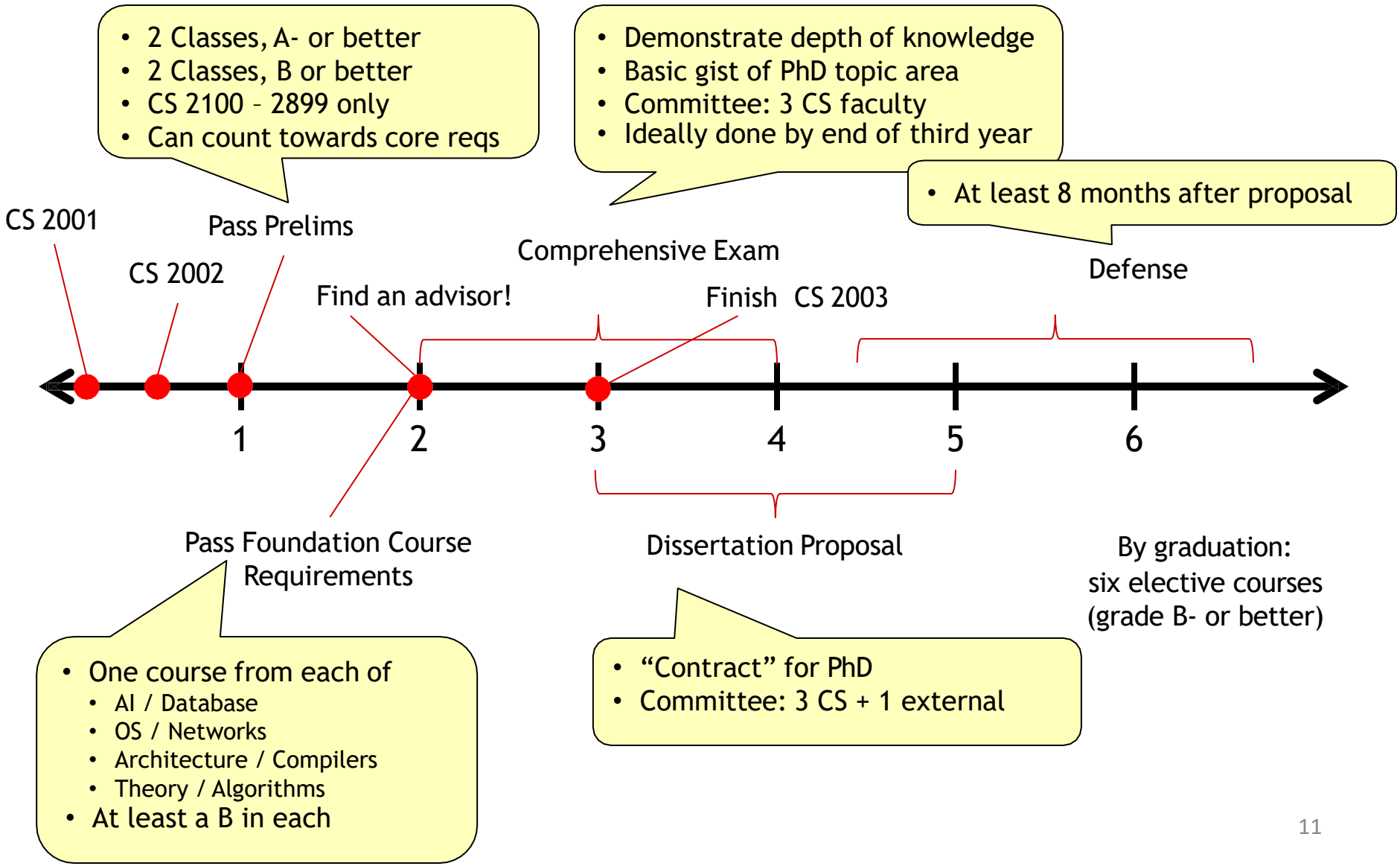
- What is your name?
- What topics within computer science are you interested in?
- What do you like to do outside of school/work?
- Why did you choose to pursue a PhD?
- Why did you choose to pursue it at Pitt?

Plan for this lecture

- PhD journey
  - PhD timeline at Pitt
  - Motivation and challenges of doing a PhD
- Managing a PhD career
  - Communicating with your advisor
  - Gaining confidence
  - Work/life balance
  - Advice from faculty
  - Publishing (discussed later)
- Research integrity
- Resources

PhD Journey

PhD Timeline at Pitt



Comprehensive exam

- "You work with three faculty members (presumably your advisor and two more) to identify one topic area per faculty paper, on which you present essentially a survey of the literature. The topic area should not be too broad (e.g. machine learning) but something a bit more specific. Depending on the area you are working in, faculty may have different expectations.
- In AI comprehensive exams, students typically choose 10-15 papers per topic, and share the list with the faculty for feedback (before the exam). The student's goal is to organize the papers in each area in some meaningful, interesting way. Show families of methods, discuss their similarities and differences, and their strengths and weaknesses. Do NOT show one slide per paper, and do not talk about any one paper in too much detail. Think of the comprehensive exam as writing a survey paper on the topic. You want to present the general directions researchers have explored, educate the audience about when they might choose one family of methods over another, and discuss limitations and open research challenges.
- In some systems comps, each topic area has 3-4 papers, but the student should know these papers really well. Sometimes faculty choose seminal papers for the exam, and sometimes faculty request students to pick the important papers in the topic area. Students are responsible for knowing the ancillary/pre-requisite materials for these 3-4 papers, and are also supposed to know how to apply the ideas of the papers to solve other problems. Not too many papers, but a lot of thinking about the papers (about the techniques and how to apply them)."

https://docs.google.com/document/d/1mcF3yHJJDdXk7PCMIn54MHufo9tIHweaVDPX5V22Ywk/

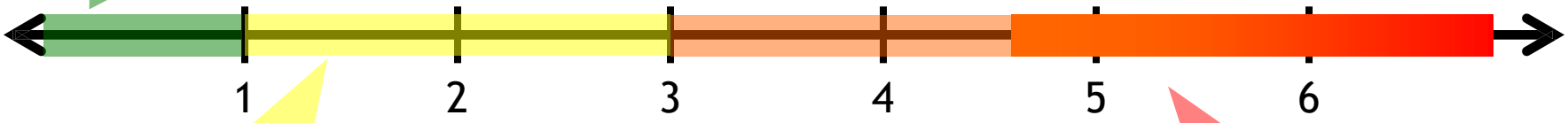
PhD Timeline at Pitt

- https://www.cs.pitt.edu/current-students/phd-computer-science/degree-requirements
  - Let's spend 5 min to read and ask questions
- https://www.cs.pitt.edu/current-students/graduate-policies
  - Let's spend 5 min to read and ask questions

PhD Timeline at Pitt

- Finish course requirements
- Identify research areas/professors of interest
- Begin reading papers, discussing area, doing pilot experiments / development

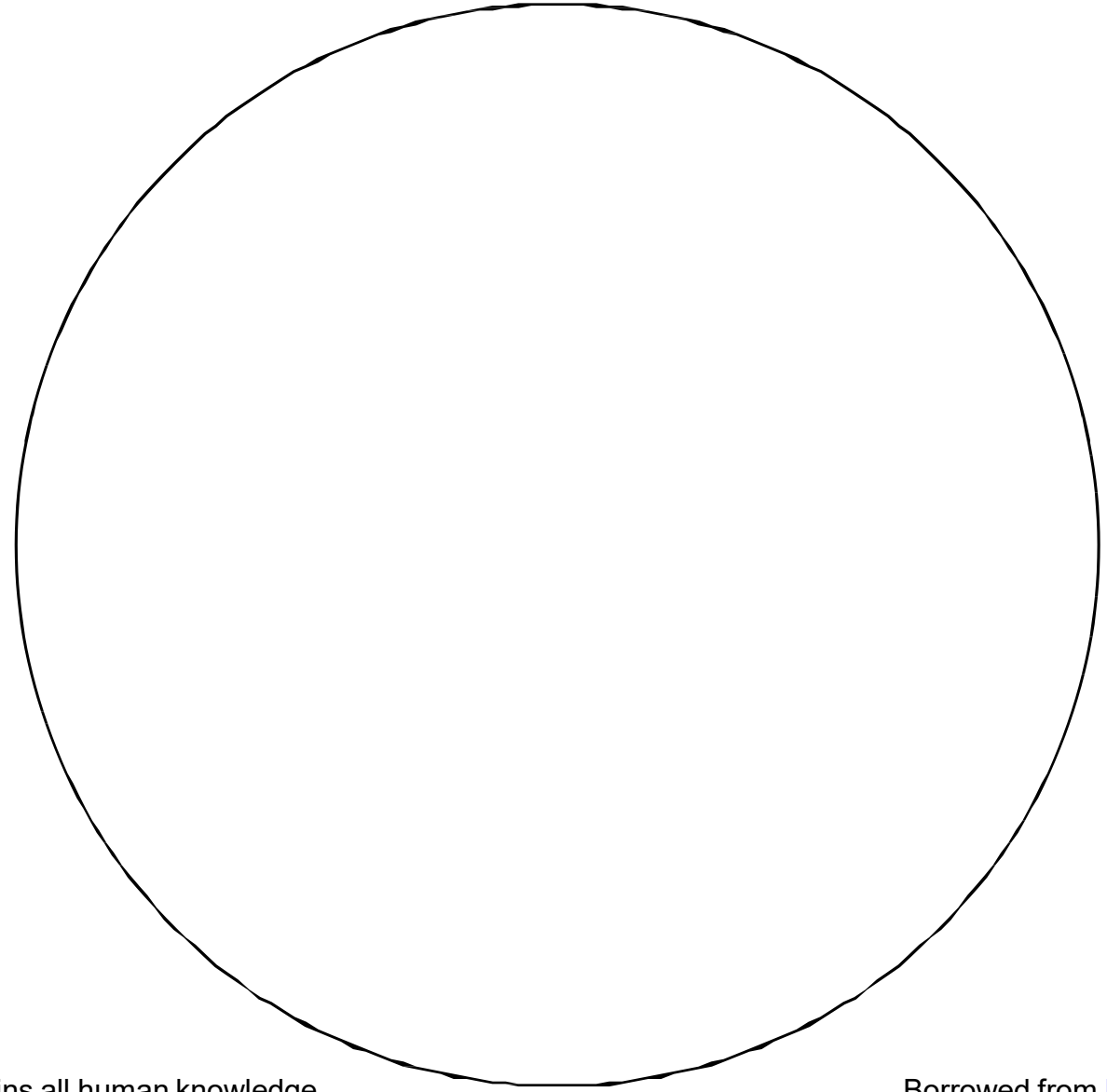
- More than smarts: Persistence!
- Time management should become an art…
- Identify larger open problem, not just many small “neat” problems
- writing, Writing, WRITING
- Continue reading



- Attend group meetings
- Read several papers per week
- Research! Develop your niche, expertise.
- Work on time management skills
- reading, Reading, READING

- Good thesis or great thesis?
- To defend, you must be an expert. Are you?
- Resist the temptation to take a job before you defend. This is a recipe for disaster.

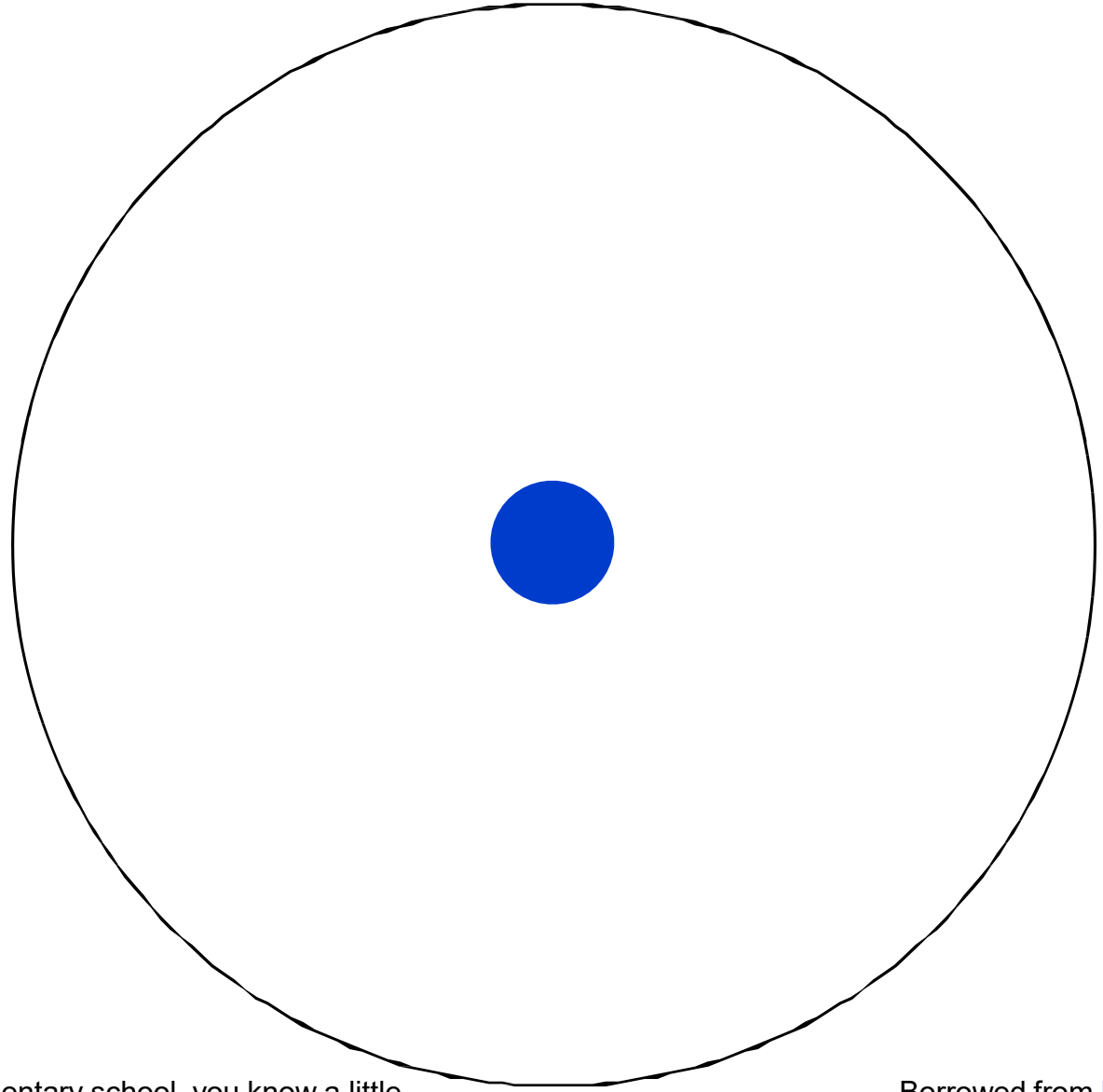
What is a PhD?



Imagine a circle that contains all human knowledge

Borrowed from http://matt.might.net

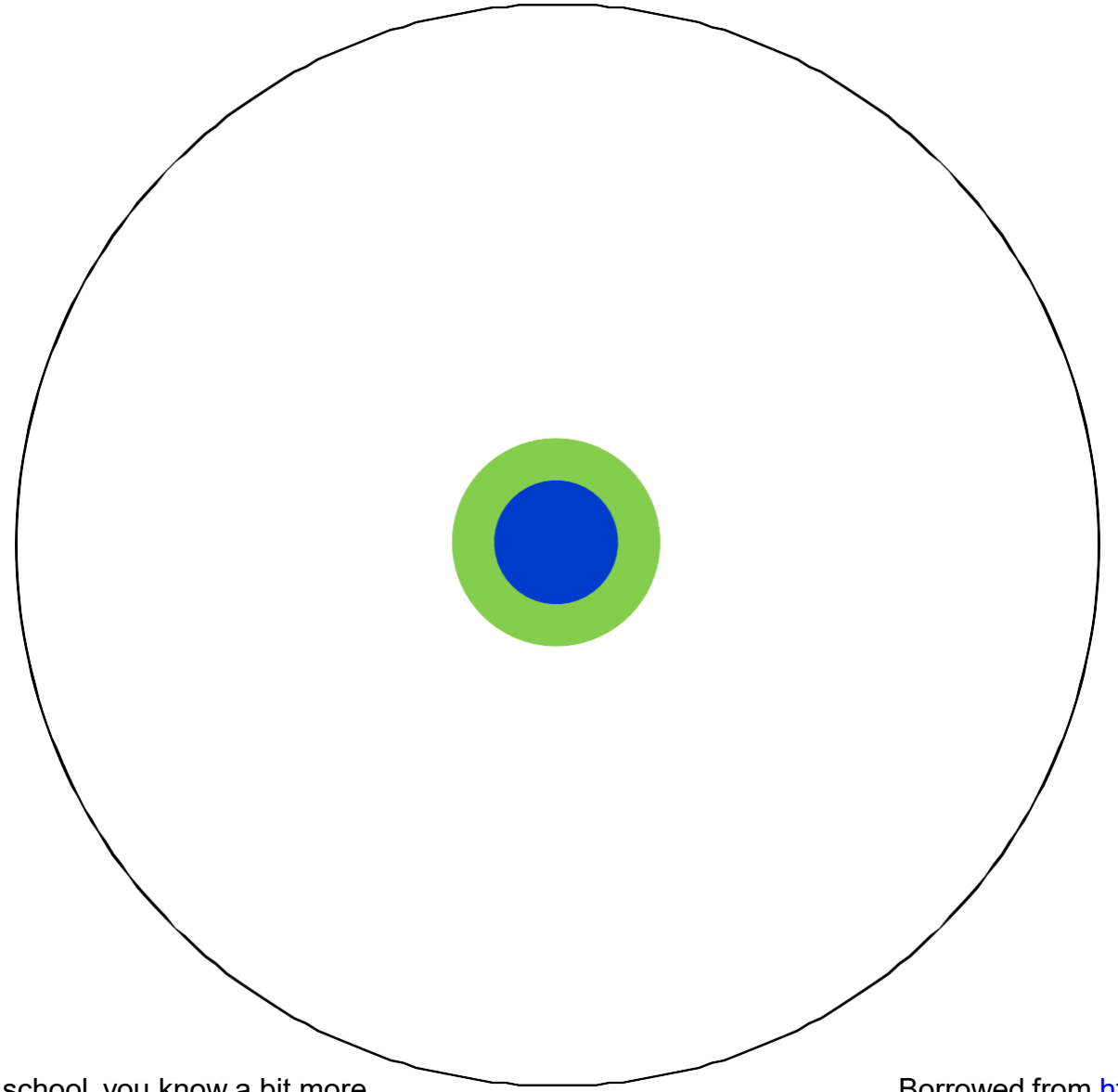
What is a PhD?



By the time you finish elementary school, you know a little

Borrowed from http://matt.might.net

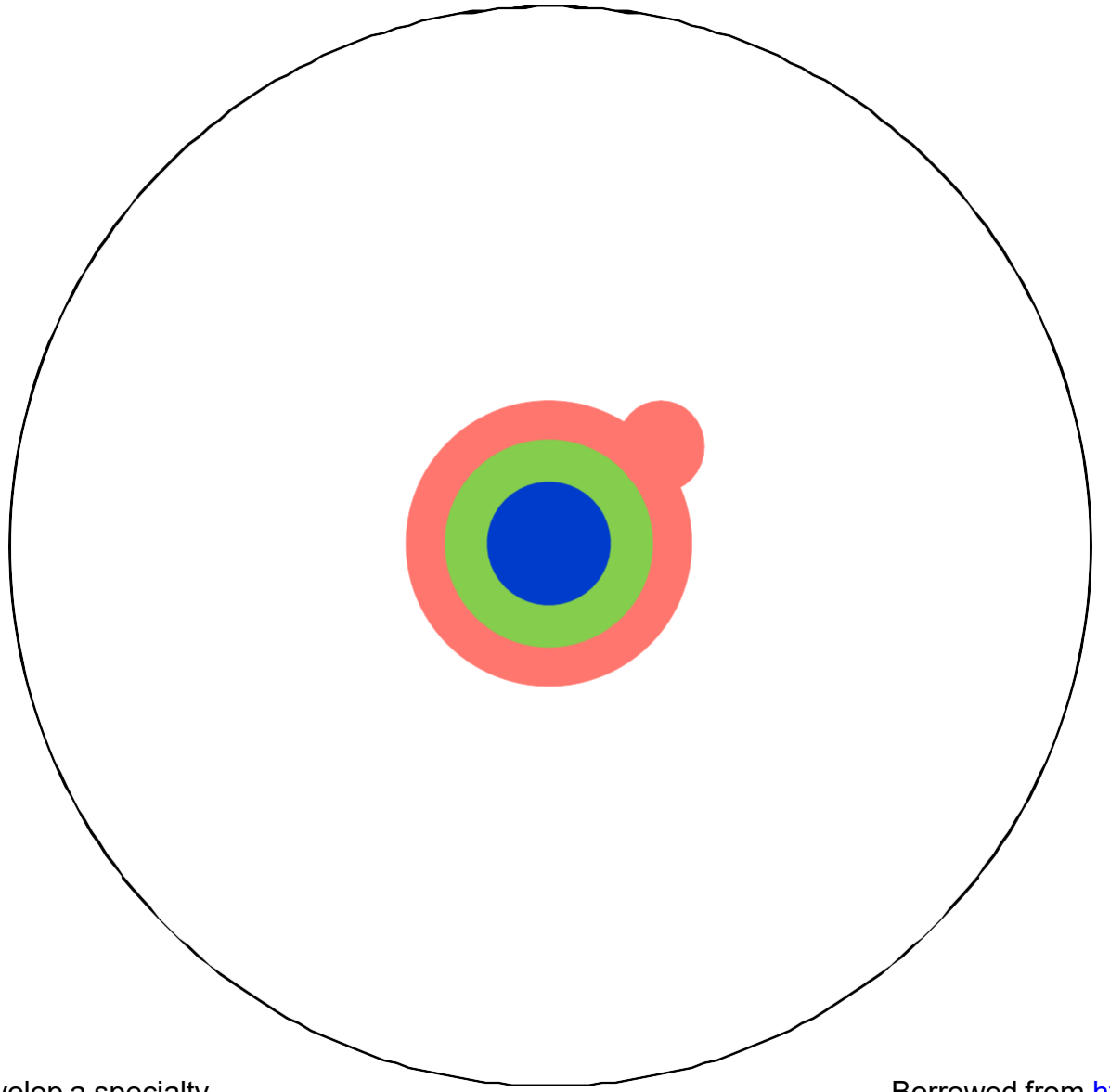
What is a PhD?



By the time you finish high school, you know a bit more

Borrowed from http://matt.might.net

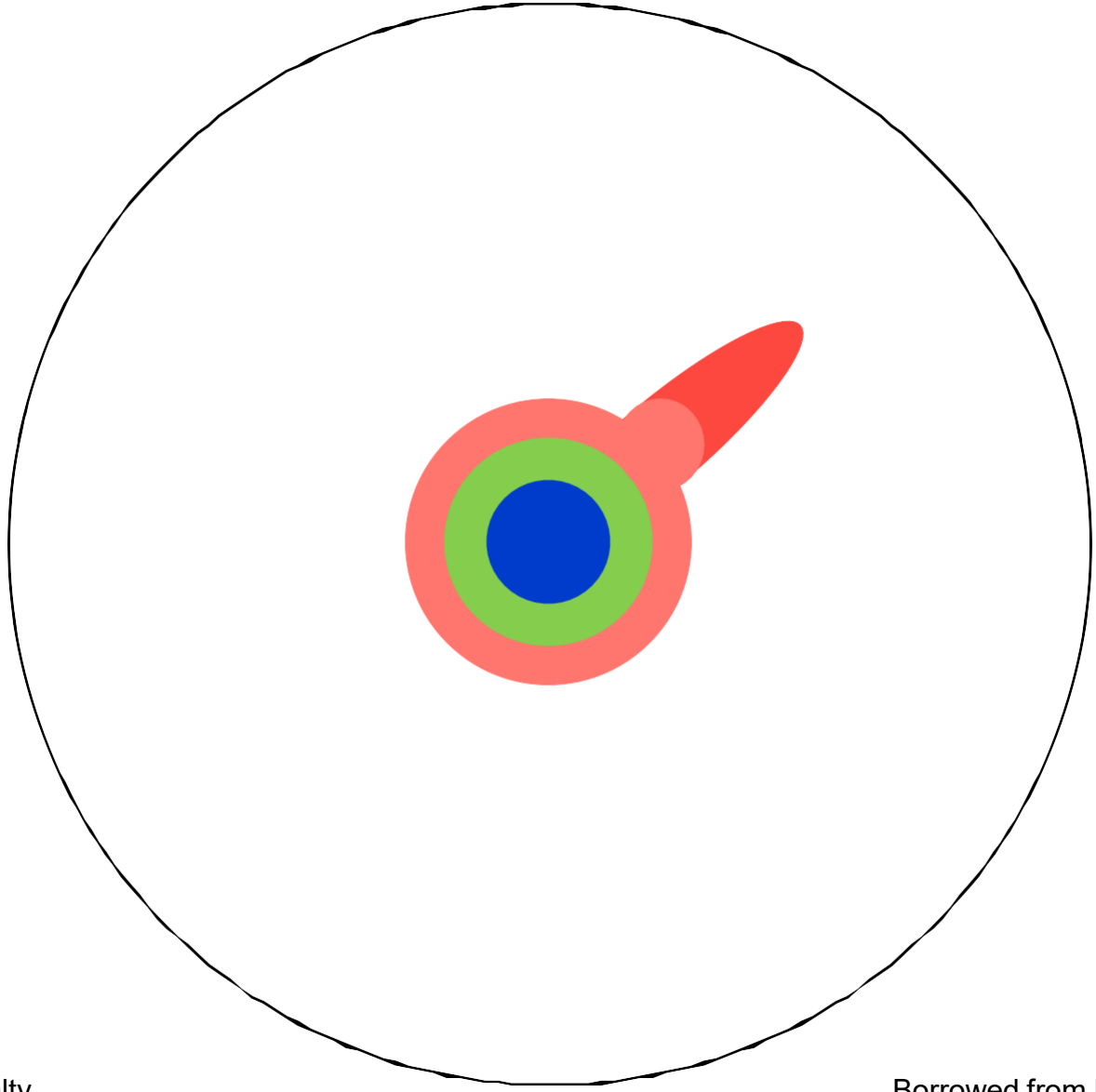
What is a PhD?



With a BS, you begin to develop a specialty

Borrowed from http://matt.might.net

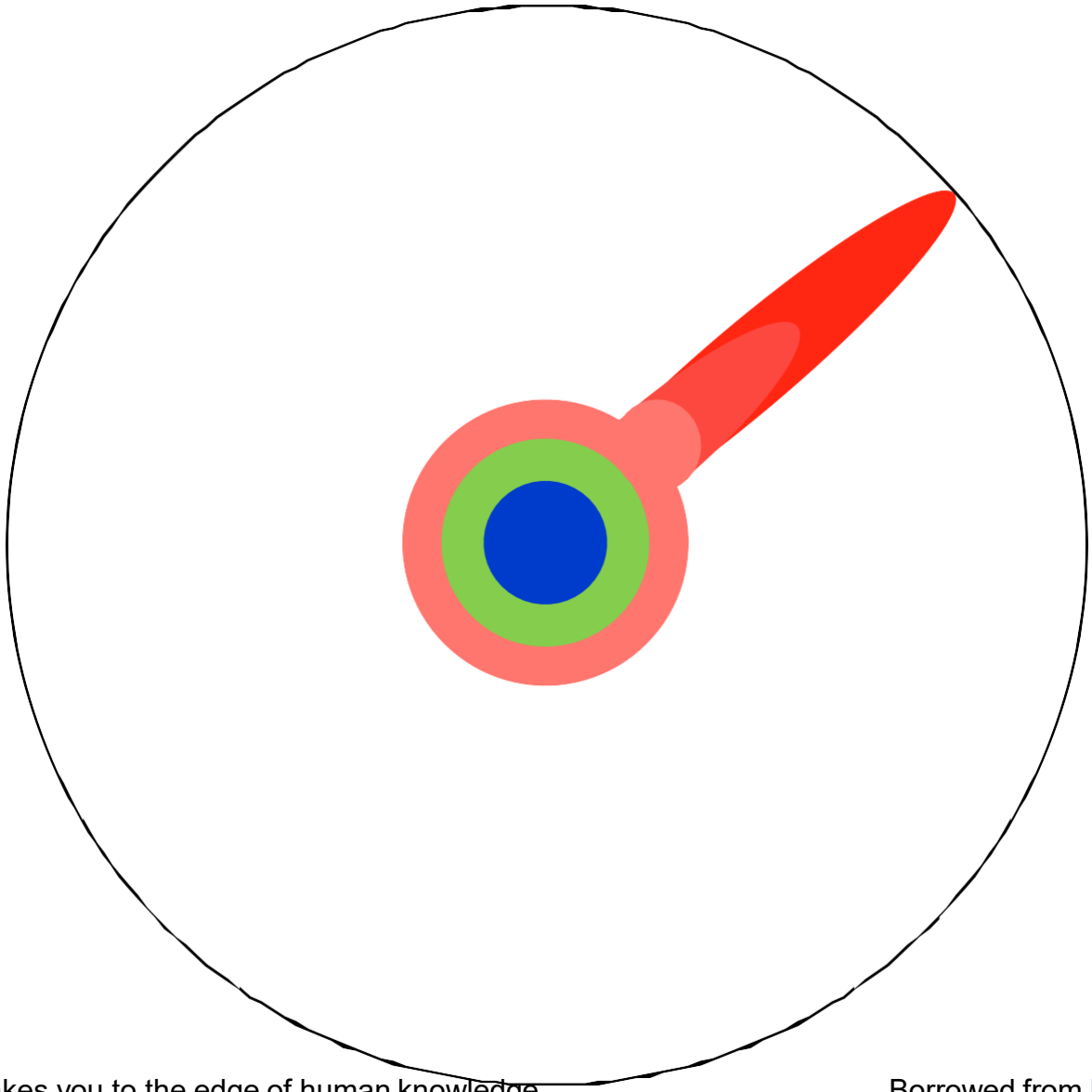
What is a PhD?



An MS deepens that specialty

Borrowed from http://matt.might.net

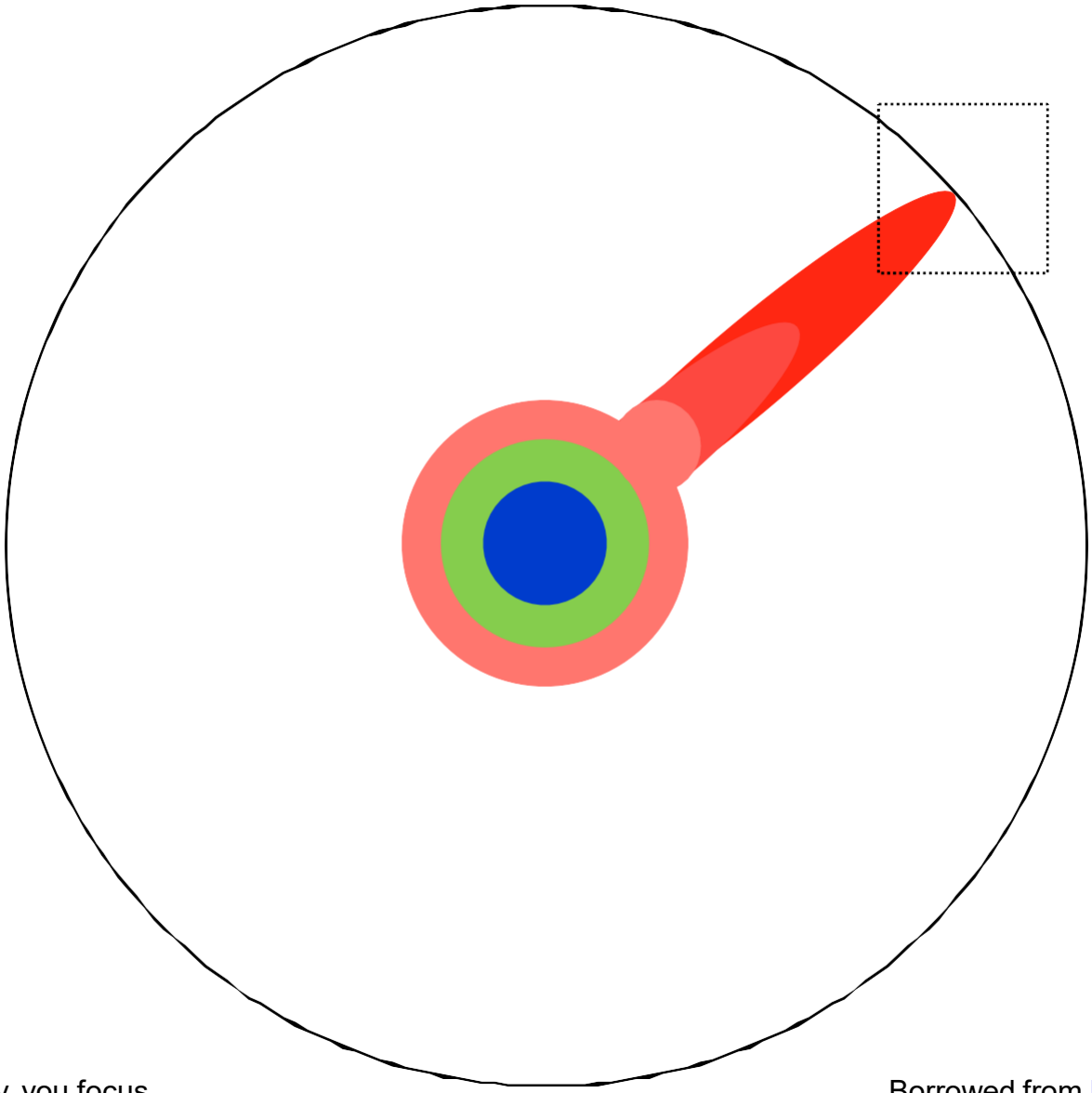
What is a PhD?



Reading research papers takes you to the edge of human knowledge

Borrowed from http://matt.might.net

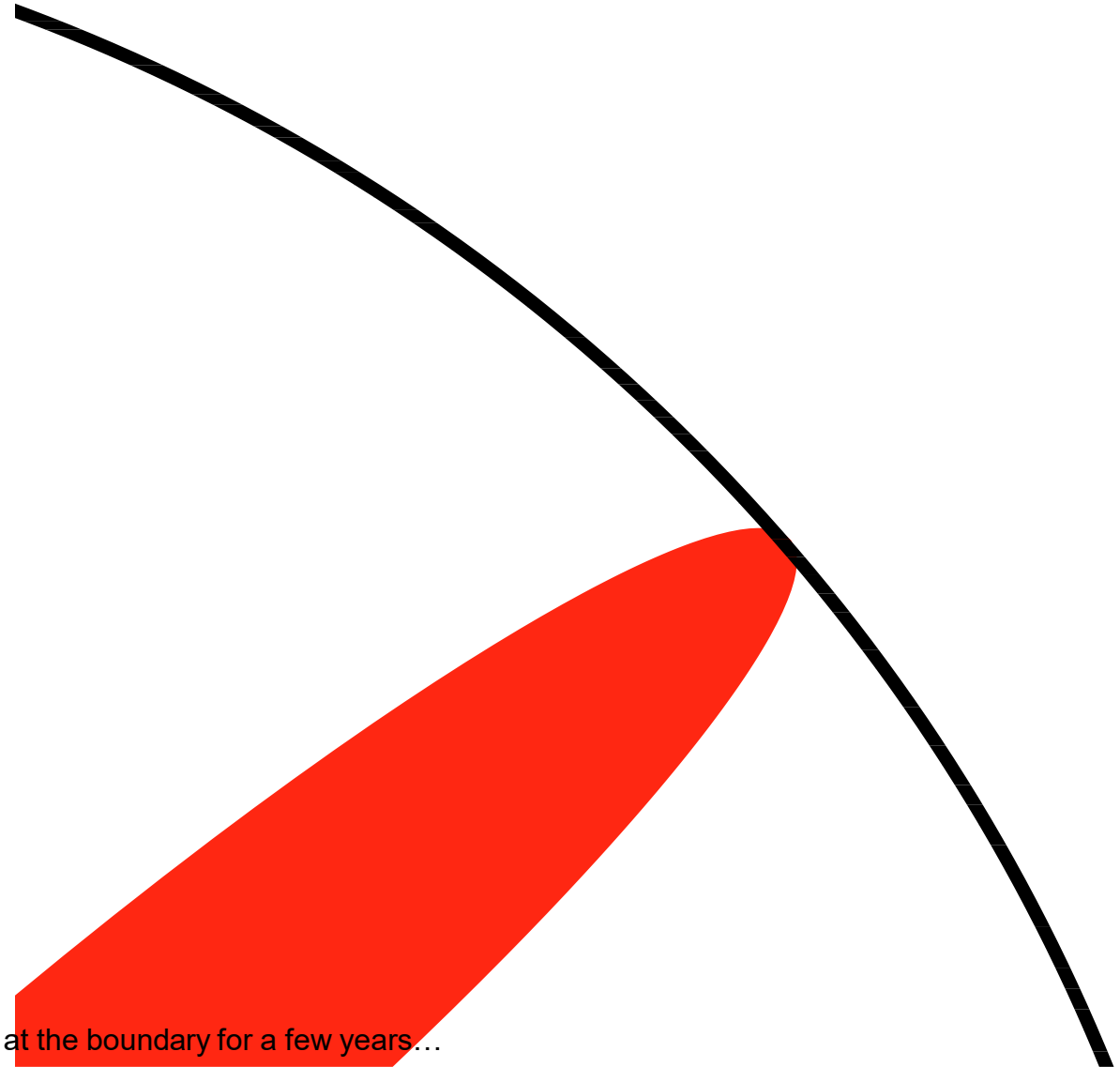
What is a PhD?



Once you're at the boundary, you focus

Borrowed from http://matt.might.net

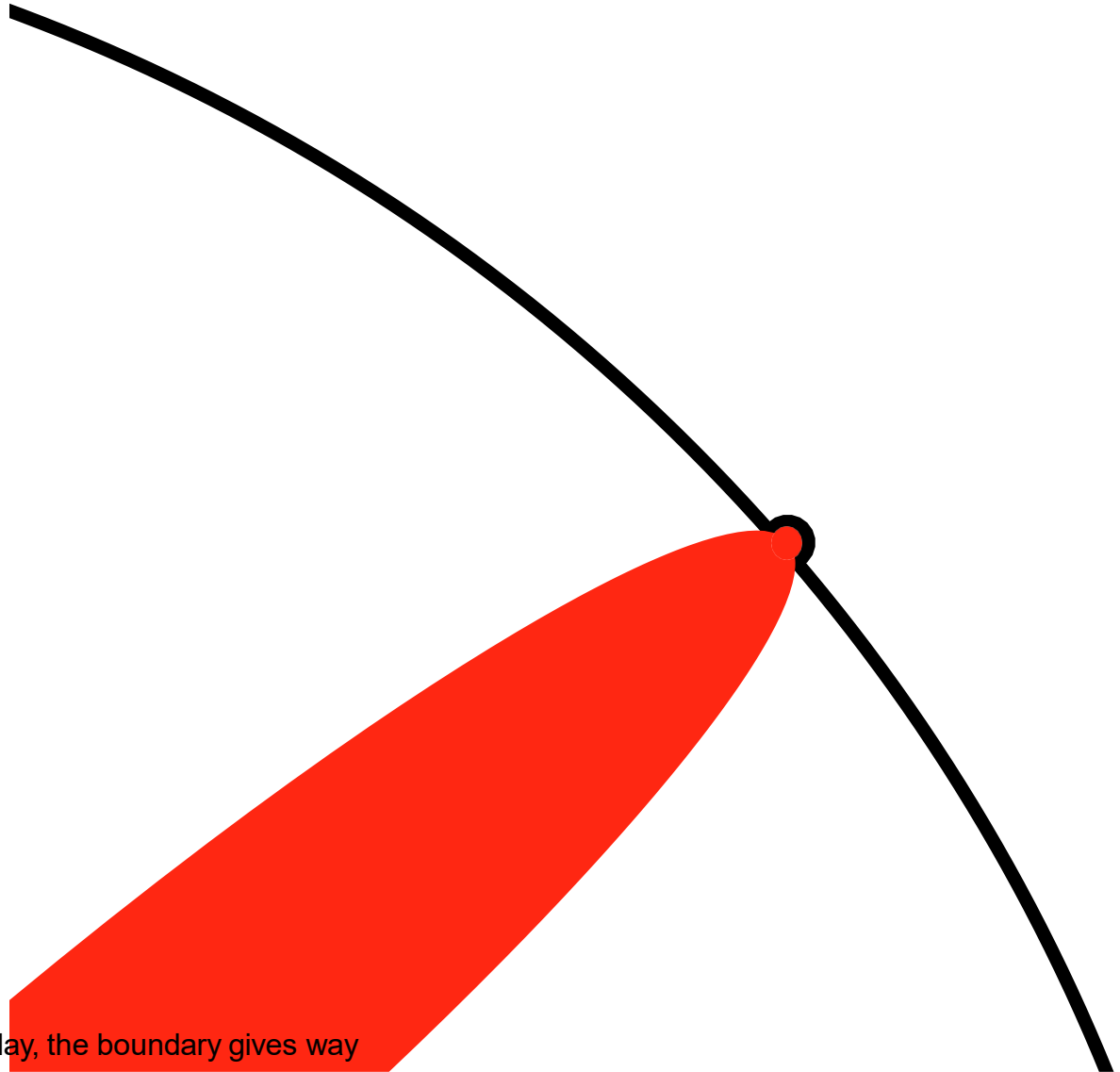
What is a PhD?



You push at the boundary for a few years…

Borrowed from http://matt.might.net

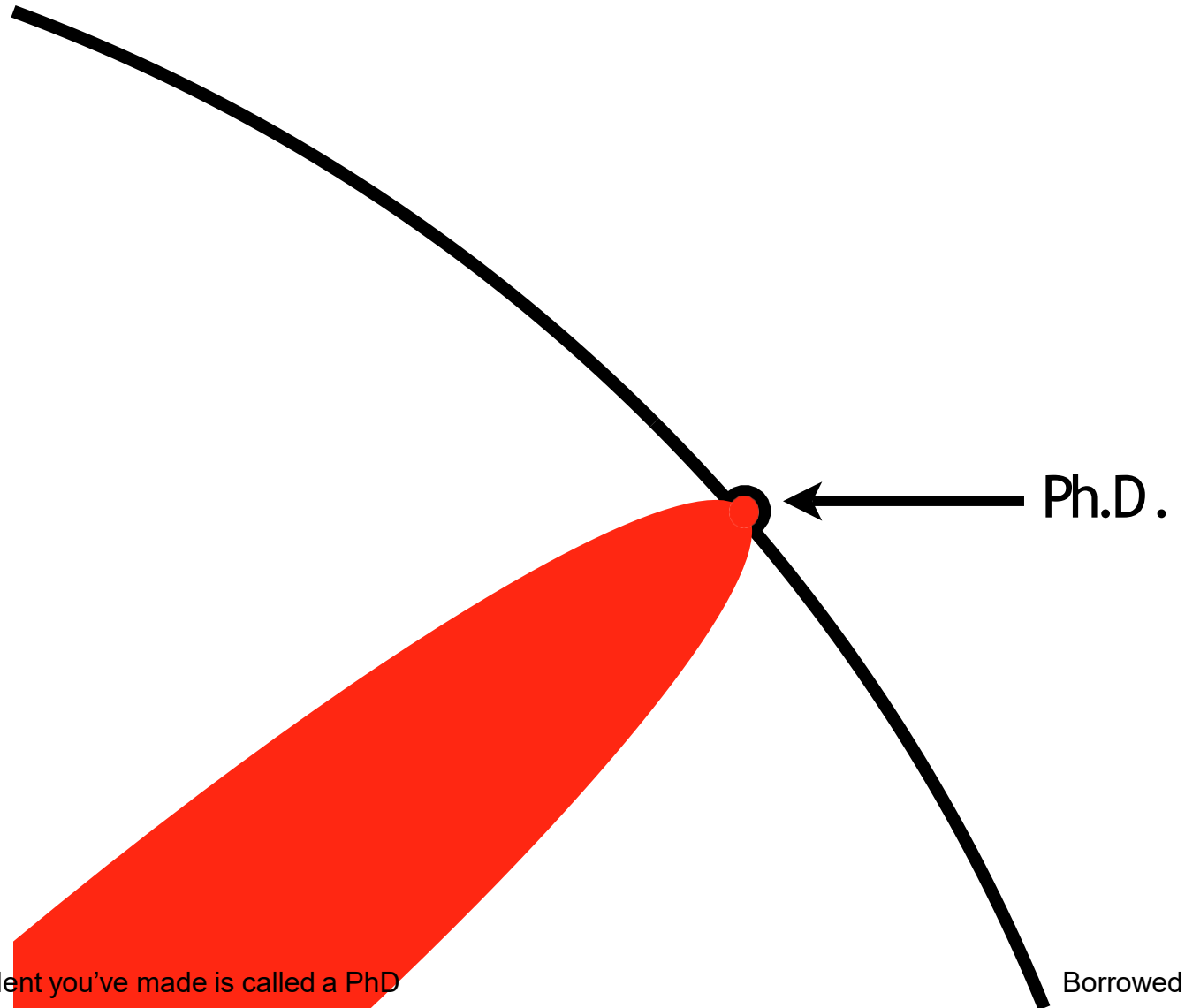
What is a PhD?



Until one day, the boundary gives way

Borrowed from http://matt.might.net

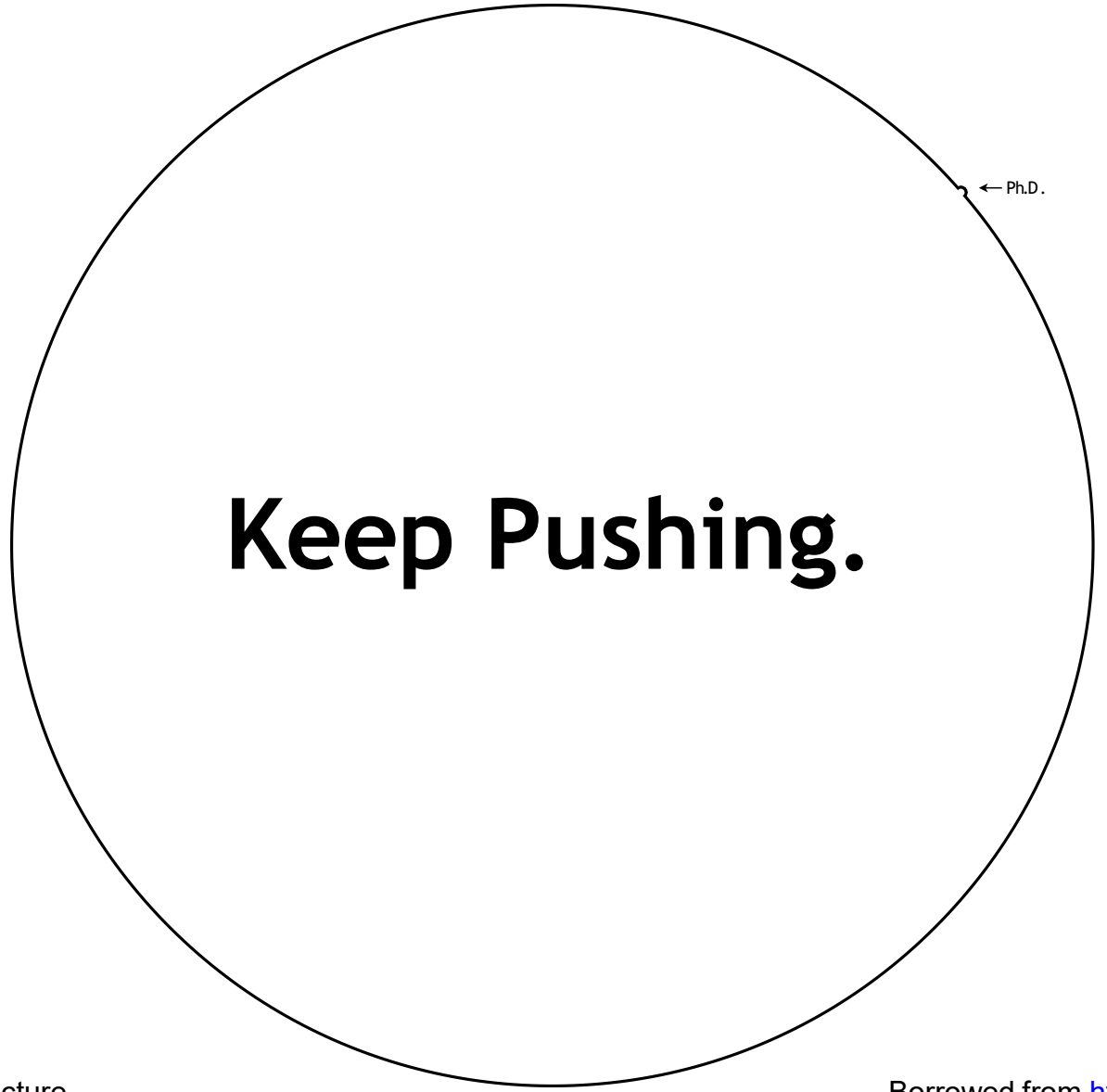
What is a PhD?



And that dent you've made is called a PhD

Borrowed from http://matt.might.net

What is a PhD?



So don't forget the bigger picture

Borrowed from http://matt.might.net

So… What *is* a PhD?

A PhD is a process!

Smart is not enough, so a demoralizing process at times…

- So much has been done already, how do I keep up?
- What new problem should I be solving?
- Solving this problem is way harder than I thought it would be…
- Why doesn't my code (or code that I've borrowed) work?!
- Why do my papers keep getting rejected?

But a process that will educate you greatly

- Problem identification and solving
- Persistence and hard work
- Clear (written and spoken) communication skills
- Team work, management, and mentoring

Slide credit: Adam Lee

My Path

- Decided to do a PhD because I thought I would enjoy learning more about AI– but no clear plan or goal
- My first paper got into CVPR with good reviews, but I quickly got demoralized because the field seemed like chasing 0.5% improvement in accuracy on a benchmark
- Then I discovered a topic I was really, really excited about, by reading a paper in reading group; my thesis ended up revolving around it
- Even so, my first attempt on this topic didn't succeed– yet a labmate later got a best paper award out of a similar idea—the line between success and failure is thin
- I learned the importance of how you present and argue the importance of your work
- I had to at least once read through related work in great detail and implement mathematical details from scratch
- I greatly enjoyed having huge time flexibility; one time, this meant I even took an acting class; more than one time, it meant spending weeks on 4 hours of sleep—but it was my choice and I felt that it was worth it

A PhD journey is long and not very well defined.
What are you worried or confused about?

Managing a PhD

Choosing an Advisor

- Choosing an advisor is a big decision…
- Key points: research interest and fit
- You are unlikely to finish a PhD that you aren't interested in…
  - Take courses and seminars in these areas
  - Read papers, make sure you're interested in recent developments
  - Think about whether you like theory vs. applications
- You will be working very closely with your advisor
  - Do you prefer "hands on" or "hands off" advising?
  - Can you take direction/criticism from this person?
  - Use CS2002 and course projects to test fit!

Slide adapted from: Adam Lee

Communicating with your Advisor

- **Your advisor** needs to know that:
  - You are independent
  - You are reliable
  - You are excited about your research
  - You will find and read literature without pointers
  - You are capable of generating (good) ideas
  - You understand the difference between low-level tweaks and mid-level technical innovation
  - You can express your ideas and motivate them

Communicating with your Advisor

- **You** need to know that:
  - (Same as above, plus)
  - Your advisor can fund you (as GSR, TA)
  - Your advisor is supportive and kind
  - Your advisor can provide the type of help you need
  - Your advisor knows about and cares enough about the direction you want to pursue
  - What else?

Meetings with your Advisor

- Likely every week, maybe more than once a week—set a schedule (more advisable), or meet as-needed
- Summarize what your goals were in the past week – what hypotheses you aimed to test, and how you tested them
- Describe the methods you read about or developed
- Describe results you obtained, and interpret them – are they reliable/significant? What do they imply? What's next?
- Describe challenges you faced, how you resolved them, and how long it took
- Your advisor may or may not want to see code, depending on research area
- Agree on goals for next week

Communication Issues: Examples

- You tried a method; you can't explain why it makes sense, or it actually doesn't make sense
- You spent a whole week working on something, but have nothing to show for it
- You waited a week for a dataset/code/something else to become available, but didn't ask your advisor for help
- You misunderstood what your advisor asked you to try

Trust and Diligence Issues: Examples

- Your code obviously has a bug (e.g. you are getting 100% prediction accuracy on a difficult AI task) but you present the results to your advisor without analyzing them
- You accidentally deleted your advisor's dataset/work
- You missed a paper deadline
- You are not responding to email during business hours, on multiple occasions

Gaining Confidence

- You have to convince yourself you can do research, starting with a small project where perhaps novelty is present, but small
- You have to be strategic about the work you do—it has to be publishable today, given the state of the field
- You have to develop a unique research identity—what do you want to be known for?

What is your work/life balance strategy?
What advice have you heard?
What advice works/doesn't?

Advice from the faculty

- Your attitude determines how far you go. Ph.D. is not about finishing the jobs I give to you, it's about solving the problems you care most and make intellectual breakthroughs. Don't expect to get a todo list from me; Instead, I will guide you to define your own research agenda.
- Prioritize research. Research is your key duty and also how the outside world will evaluate you when you leave here. Learning how to de-prioritize other things without annoying other people or hurting yourself is a skill.
- Focus on building foundational knowledge for your long-term career. Abraham Lincoln: "Give me six hours to chop down a tree and I will spend the first four sharpening the axe."
- Have long-term and short-term plans for your research.
- Get your hand dirty.
- Document code and record experimental settings and how datasets were filtered.

https://docs.google.com/document/d/1mcF3yHJJDdXk7PCMIn54MHufo9tIHweaVDPX5V22Ywk/

Advice from the faculty (cont'd)

- Read lots of papers: keep up-to-date with developments in your field. Focus on understanding their contributions, problem statement (what problem are they addressing), novelty (what new perspective or new method are they introducing), etc. Reading paper guide: http://ccr.sigcomm.org/online/files/p83-keshavA.pdf
- Be skeptical when reading papers. Ask yourself: does this research problem make sense to me? Is there a simple way to solve this problem? Why conduct the experiment this way, not that way?
- Be curious about other people's work, even if it is not "your topic/area" (you can learn from listening and even more from asking questions and giving feedback when appropriate)
- Journal your ideas and discuss them with peers - discuss your ideas with PhD students to get different perspectives.

https://docs.google.com/document/d/1mcF3yHJJDdXk7PCMIn54MHufo9tIHweaVDPX5V22Ywk/

Advice from the faculty (cont'd)

- Focus on progress, not perfection. Have something to show your advisor when you meet even if it's not perfect, or provide reasons for why you weren't able to make progress. For papers, try to write a rough draft early even if it's very poor. […] The process of writing and presenting research often reveals areas where you might need to do some more coding or analysis. It's ideal to find those early.
- Try to "pop up a level" to describe what you're doing and why relatively frequently. That can help you avoid going down rabbit holes of debugging code forever or building small apps or endlessly tweaking datasets, etc, where there might be quicker and easier alternatives. Writing down weekly goals and accomplishments and discussing those with your advisor can help retain your focus on what matters.
- In your research meeting with your advisor, think ahead and be prepared. Prepare to present your results, answer questions that your advisor might ask you, and propose the next steps.

https://docs.google.com/document/d/1mcF3yHJJDdXk7PCMIn54MHufo9tIHweaVDPX5V22Ywk/

Advice from the faculty (cont'd)

- Be a nice person. Academia is full of smart people, but being a nice person could help you go a long way.
- Build a cohort of fellow doc students early. Lean on them for peer-mentorship, but also friendship. Otherwise, the PhD can be lonely journey.
- Follow your heart and focus on things you feel most interesting and important

https://docs.google.com/document/d/1mcF3yHJJDdXk7PCMIn54MHufo9tIHweaVDPX5V22Ywk/

Research Integrity

Ethical Research

- Plagiarism—a big problem!
  - Don't claim others' work as your own
  - Don't copy text from others' work—ideally use quotes, but preferably say with your own words
- Verify your code and evaluation is sound!
  - Proofread code often, ask peers to review
  - Don't fabricate results!
- IRB human subjects research
  - Talk to your advisor

Ethical Research (cont'd)

- Keep in mind social impact
  - Who does my research help, hurt?
  - Examples: Generative AI—help/hurt?
  - Facial recognition software—help/hurt?
  - Software that can predict speaker's face from their voice—help/hurt?
  - Predicting political bias from politicians' portrayals (study how Obama's face may appear in left/right-leaning media)—help/hurt?

Research Integrity

- "**Research misconduct** is defined as fabrication, falsification, or plagiarism, including misrepresentation of credentials, in proposing, performing, or reviewing research, or in reporting research results. It does not include honest error or differences of opinion. Misconduct as defined above is viewed as a serious professional deviation that is subject to sanctions imposed both by the University, by many professional associations, and, in the case of research proposed to or funded by a federal agency, by that agency."

http://www.provost.pitt.edu/documents/GUIDELINES%20FOR%20ETHICAL%20PRACTICES%20IN%20RESEARCH-FINALrevised2-March%202011.pdf

Research Integrity (cont'd)

- "A researcher **must not report anticipated research results that had not yet been observed** at the time of submission of the report. In order to preserve accurate documentation of observed facts with which later reports or conclusions can be compared, every researcher has an obligation to **maintain a clear and complete record of data acquired**. As stated in the University's Guidelines on Research Data Management, "records should include sufficient detail to permit examination for the purpose of replicating the research, responding to questions that may result from unintentional error or misinterpretation, establishing authenticity of the records, and confirming the validity of the conclusions."

http://www.provost.pitt.edu/documents/GUIDELINES%20FOR%20ETHICAL%20PRACTICES%20IN%20RESEARCH-FINALrevised2-March%202011.pdf

Research Integrity (cont'd)

- "Authors who present the words, data, or ideas of others with the implication that they are their own, without attribution in a form appropriate for the medium of presentation, are committing **theft of intellectual property and may be guilty of plagiarism and thus of research misconduct**. This statement applies to reviews and to methodological and background/historical sections of research papers as well as to original research results or interpretations. If there is a word-for-word copying beyond a short phrase or six or seven words of someone else's text, that section should be enclosed in quotation marks or indented and referenced, at the location in the manuscript of the copied material, to the original source. The same rules apply to […] student papers submitted for academic credit. […] it could constitute an infraction of the law by infringing on a copyright held by the original author or publisher. An author should cite the work of others even if he or she had been a co-author or editor of the work to be cited [...]"

http://www.provost.pitt.edu/documents/GUIDELINES%20FOR%20ETHICAL%20PRACTICES%20IN%20RESEARCH-FINALrevised2-March%202011.pdf

Research Integrity (cont'd)

- "One particularly serious form of plagiarism is the **misuse of privileged information taken from a grant application or manuscript received from a funding agency or journal editor for peer review**. In such a case, the plagiarism is a serious matter of theft of intellectual property because it not only deprives the original author of appropriate credit by citation but could also preempt priority of first publication or use of the original idea to which the source author is entitled. Also, one who breaches confidentiality by showing a privileged unpublished document to an unauthorized person can be held to a shared responsibility for any subsequent plagiarism of the document committed by that unauthorized person."

http://www.provost.pitt.edu/documents/GUIDELINES%20FOR%20ETHICAL%20PRACTICES%20IN%20RESEARCH-FINALrevised2-March%202011.pdf

Research Integrity (cont'd)

- "Research integrity requires not only that reported conclusions are based on accurately recorded data or observations but that **all relevant observations are reported. It is considered a breach of research integrity to fail to report data that contradict or merely fail to support the reported conclusions**, including the purposeful withholding of information about confounding factors. If some data should be disregarded for a stated reason, confirmed by an approved statistical test for neglecting outliers, the reason should be stated in the published accounts. A large background of negative results must be reported. Any intentional or reckless disregard for the truth in reporting observations may be considered to be an act of research misconduct."

http://www.provost.pitt.edu/documents/GUIDELINES%20FOR%20ETHICAL%20PRACTICES%20IN%20RESEARCH-FINALrevised2-March%202011.pdf

Research Integrity (cont'd)

- "Research data obtained in studies performed at the University of Pittsburgh and/or by employees of the University are **not the property of the researcher** who generated or observed them or even of the principal investigator of the research group. They belong to the University of Pittsburgh, which can be held accountable for the integrity of the data even if the researchers have left the University."
- "Each student, postdoctoral fellow, or other investigator in a group project should come to an **understanding with the research director or principal investigator**, preferably in writing, about which parts of the project he or she might continue to explore after leaving the research group."

http://www.provost.pitt.edu/documents/GUIDELINES%20FOR%20ETHICAL%20PRACTICES%20IN%20RESEARCH-FINALrevised2-March%202011.pdf

Research Integrity (cont'd)

- "Since the scientific enterprise may be a cooperative endeavor encompassing many persons who now or in the future might pursue related research interests, and since it is in the interest of all to rely on the contributions and findings of others, every investigator has an obligation to the general scientific community to cooperate by **sharing of data.** Other virtues of sharing data include the facilitation of independent confirmation or refutation of reported outcomes. It is generally accepted that the data underlying a research publication should be made available to other responsible investigators upon request after the research results have been published or accepted for publication."

http://www.provost.pitt.edu/documents/GUIDELINES%20FOR%20ETHICAL%20PRACTICES%20IN%20RESEARCH-FINALrevised2-March%202011.pdf

Research Integrity (cont'd)

- "Publication must give appropriate **credit to all authors** for their roles in the research. If more than one person contributes significantly, the decision of which names are to be listed as co-authors should reflect the relative contributions of various participants in the research. Many professional associations and research journals have specified criteria for authorship. One common standard appearing in many of these statements is that each author should have participated in formulating the research problem, interpreting the results, and writing the research paper, and should be prepared to defend the publication against criticisms. […] A person's name should not be listed as author without his or her knowledge, permission, and review of the final version of the manuscript that includes the names of all co-authors."

http://www.provost.pitt.edu/documents/GUIDELINES%20FOR%20ETHICAL%20PRACTICES%20IN%20RESEARCH-FINALrevised2-March%202011.pdf

Research Integrity (cont'd)

- "Research protocols involving **human subjects** must be approved in advance by the University Institutional Review Board (IRB), which determines whether risks posed to subjects are acceptable and whether information describing risks and benefits of subject participation is conveyed to subjects in accurate & intelligible manner."
- "Research means '...a systematic investigation, including research development, testing and evaluation, designed to develop or contribute to generalizable knowledge.' Human subject means '...a living individual about whom an investigator (whether professional or student) conducting research obtains (1) data through intervention or interaction with the individual, or (2) identifiable private information' "
- "There may be a number of scholarly activities (e.g., an oral history project) which may not meet the federal definition of research, or other activities (e.g., secondary data analyses using de-identified data) which may be meet the federal definition of research but may not meet the federal criteria for the involvement of 'human subjects.'

http://www.provost.pitt.edu/documents/GUIDELINES%20FOR%20ETHICAL%20PRACTICES%20IN%20RESEARCH-FINALrevised2-March%202011.pdf

Resources

Computing etc. Resources

- Department labs
- Department cluster (RIC)
- Department storage (AFS)
- Your advisor's machines
- Center for Research Computing (CRC)
- Pittsburgh Supercomputing Center (PSC)
- Pittsburgh Quantum Institute (PQI)
- Google's Colab
- Code sharing and storage: GitHub
- Collaborative writing: Overleaf
- Reading: Google Scholar, DBLP, etc

Research Fellowships

- Pitt: CS50
- US government: NSF Graduate Research Fellowship
- Industry: Google, Facebook, Microsoft, Amazon, IBM, NVIDIA, Qualcomm, etc.
- Foundations: Anita Borg
- Slightly outdated list: https://docs.google.com/spreadsheets/d/1UpZlMvZ4AMK41MDa8zlnS817BJeeK8ZoCFdhs_AaVbQ/

Writing Resources

- The Writing Center: https://www.writingcenter.pitt.edu/

Mental Health Resources

- University Counseling Center: https://www.studentaffairs.pitt.edu/cc/

Welcome to the program!

Questions? Comments? Concerns?