Introduction to Research

(1) Creating Impact

Rex Ying

Research is about Impact

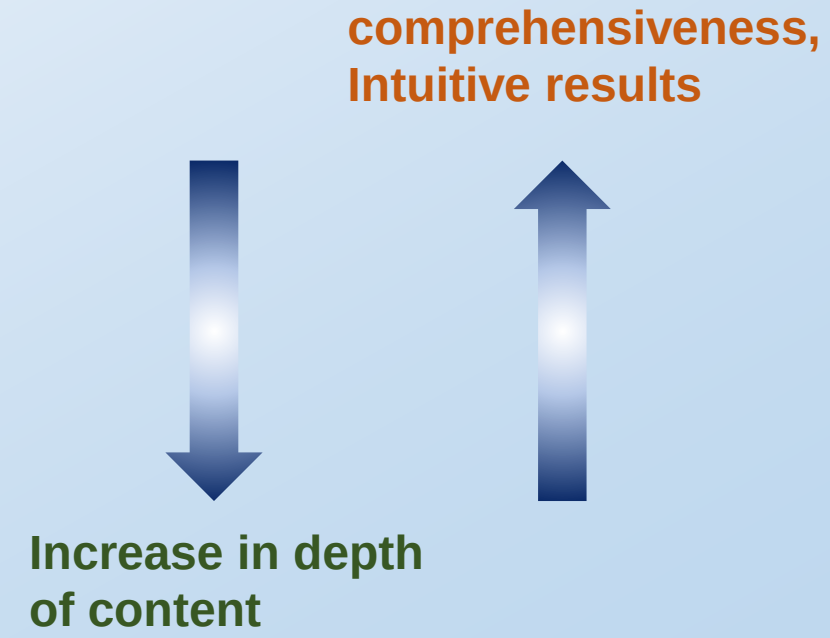
- Central question in your PhD study:
  - How does your research benefit the **research community** and **beyond**?
- I'm here to help you achieve this goal
  - Industry or academia jobs?
  - Effective communication
  - Direct feedback

Publications

- Direct Impact through **Research papers and publication**
  - We'll have a separate discussion on how to make publications impactful

- **Publication** is not the whole story

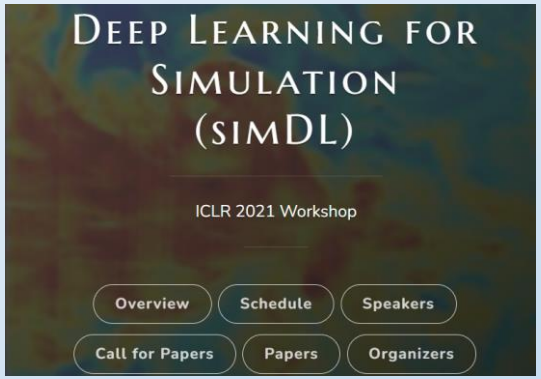
- Be aware of the types of audience
  - General public, students
  - Researchers / engineers with CS background
  - AI / ML practitioners and researchers
  - Experts in your field



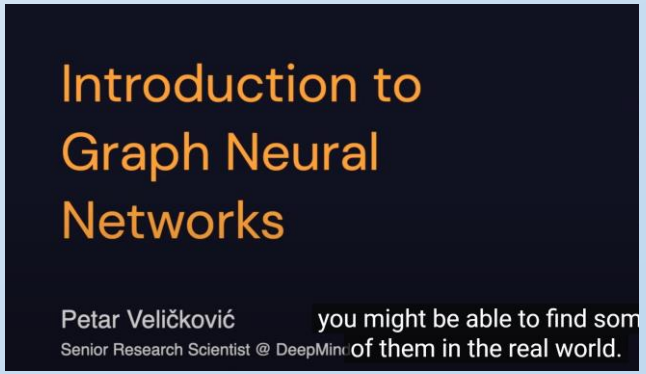
Presentations and Talks

- Talks are great ways to **promote your research**, share your findings and let people be more aware of its significance
- Venues:
  - Conferences, tutorials, workshops
  - Other academic institutions, industry collaborators, grant meetings
- Do not limit yourselves to only ML conferences!!

Conferences



Workshops



Research talks

Teaching and Mentorship

- If you want to join academia, **learning to teach** is crucial
- Be a teaching assistant
  - Make pedagogical **slides** that are easy to understand
  - **Homework** / exam questions
  - **Office hours**: answer questions but do not directly give away answers
  - Students might be inspired by your research or apply your research in their future jobs!
- Be a research mentor (when you are more senior in research)
  - Take research assistantship / independent study students interested in research
- Organize **reading groups** on research topics

Impact: Tool-building

- Beyond talks, there are other opportunities to promote your research
- Be proactive in identifying such opportunities
- Tool building is a great way to accelerate research in the field
  - **Research codebase** with high coding standards
  - **Libraries** on a specific research field
  - **Benchmarks** and evaluation frameworks
  - These tools will also make it easy when applying your research to real-world use cases

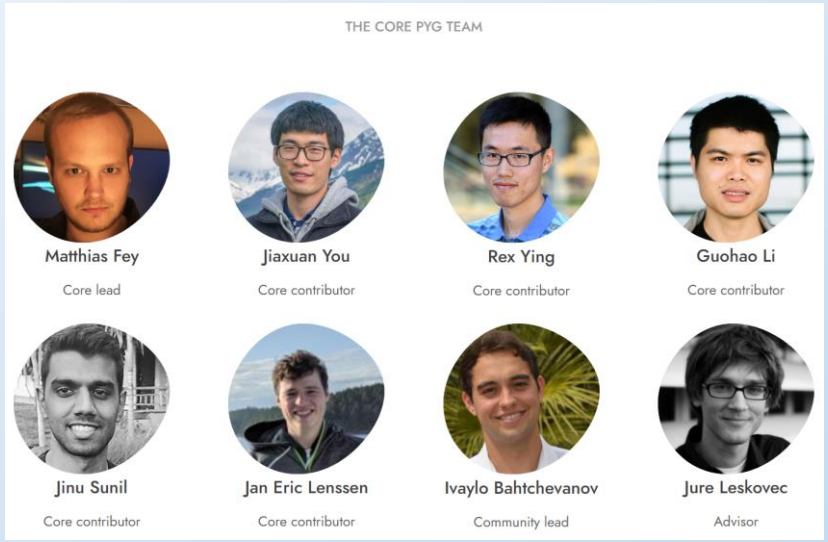
High Quality Codebase

- High-quality code is crucial for reproducibility in research
- Increase **impact of research** by allowing different audience to use your code
- Principle
  - **Simple API**! With very few lines of code, the user should be able to load in their own data, train a model, and make inference / prediction
  - **Modular**: make sure different functionalities are compartmentalized. Use OOP (Python allows multiple inheritance - Mixins!)
- More about coding and engineering practices in a separate session

This is typically a starting point for great libraries and tools!

Libraries

- Learning framework for a class of models: PyG / DGL



- Framework targeting use cases: KGE,
- Specific sub-topic: PyG-temporal, GNN AutoML
- There is high variation

APIs, Engineering designs, Efficiency, Coding styles and Active maintenance are all crucial factors!!

Benchmarks and Evaluation Frameworks

- Examples
  - OGB: large scale graph learning benchmarks
  - GraphFramEx: graph explainability evaluation and benchmarks
  - DawnBench: training and inference speed / performance benchmarks
  - GraphWorld: synthetic graphs with diverse structure
  - Knowledge graphs, molecules, proteins, physical simulations, graph generative models …
- Again, be **proactive** in finding a **unique angle** to evaluation and benchmarks

Competitions

- Organize or participate in competitions

- KDD Cup Competitions
  - Graph AutoML
  - OGB large-scale challenge
  - Focus on a specific area / challenge that with real-world significance
- NeurIPS Competition Track
- Other venues in science and technology

Collaborations

- Impact beyond machine learning / CS: **interdisciplinary research**
  - Many professors are willing to explore ML methods in their research
  - Talk to people from different backgrounds at the university, conferences or other venues
- **Industrial internships** are very important!
  - Understand real-world challenges
  - Get access to data that are very different from typical ML benchmarks
  - Create collaboration between the lab and the company
  - Get recommendation letter!
- **Grants** with other universities / industrial partners

PhD Fellowships

- There are many opportunities to obtain a **PhD fellowship**
  - They provide financial support
  - some may require / suggest internships
  - Google, Apple, Nvidia, Meta etc. all provide fellowship opportunities
    - Search online for eligibility. Be prepared when eligible.
  - Some of them may require **nomination** from the department
- **Benefits**
  - Great to appear in your CV!
  - Establish bonds with industry

Constantly look for such opportunities

Summary

- Impact is a multi-dimensional goal
  - **Research**: publications, conference presentations, workshops, tutorials, talks
  - **Teaching**: TA, mentoring, reading groups
  - **Tool-building**: high-quality code, libraries, benchmarks, competitions
  - **Collaborations** (academia, industry, cross-disciplinary)
  - **Fellowships and awards**

- Communication is crucial
  - Talk to me when you need help

- Be **proactive**!!

More About Research and PhD

1. **Creating Impact**
2. Honor Code and Research Environment
3. Collaborations
4. Research and Paper Writing
5. How to Create Great Figures
6. Coding and Engineering Practices
7. Services