

**Distributed DNN Training Grant Proposal**  
**Zhengqi Dong, Rayan Hamza, Nawras Al Naasan, and Arpan Jain**

**2.2.2 Learn about the SBIR Program and NSF CSSI and CDS&E Programs**

**NSF CSSI Program**

- **What are their foci?**
  - They focus on Institutions of Higher Education (IHEs), or Non-profit, non-academic organizations, or NSF-sponsored federally funded research and development centers (FFRDCs) projects that related to cyberinfrastructure (CI)
- **What kinds of projects do they typically fund?**
  - They fund the project that meet the vision and blueprint demonstrated in this document, <https://www.nsf.gov/cise/oac/vision/blueprint-2019/>, and based on software and data that are findable, accessible, interoperable, reusable, provenance traceable, and sustainable, and the project must designed to overcome significant bottlenecks to solving compelling S&E questions
- **What might it take to be successful in obtaining a grant?**
  - a) Follows the needs in cyberinfrastructure (CI):
    - i) Identify science and engineering challenges where the proposed CI services enable fundamental new science and engineering advances, and describe how the proposed project fosters partnerships and community development that will have a significant impact on science and engineering research;
    - ii) Indicate how the proposed CI services build capability, capacity and cohesiveness of a national CI ecosystem;
    - iii) Clearly articulate the delivery and outreach mechanism with quantifiable targets for metrics to measure impact; and
    - iv) Provide a compelling discussion of the CI's potential use by a wider audience and its contribution to a national CI.
    - v) (See Section VI.A.2 for details on the merit review criteria.)
  - b) High quality of proposal that match the vision demonstrated summarized in the [Transforming Science Through Cyberinfrastructure: NSF's Blueprint for a National Cyberinfrastructure Ecosystem for Science and Engineering in the 21st Century](#)
- **How do they align with your current project? How might you have to adjust your work to meet their foci?**
  - Our project focuses on training out-of-core deep learning models for large images, which accord with the need of “increasing the availability and scales of computation and data”, -- <https://www.nsf.gov/cise/oac/vision/blueprint-2019/>.
- **Funding availability:**
  - *Elements* awards shall not exceed a total of \$600,000 and 3 years duration (up to \$200,000 per year).
  - *Framework Implementations* awards shall range from \$200,000 to \$1M per year and shall be 3 to 5 years in duration.

**NSF HDR Program**

- Their focus is typically on academic institutions such as Universities, many of their past recipients were graduate students or academic faculty from various universities in the United States.
- The HDR program seeks to fund those who are constructing frameworks for data-intensive research in science and engineering.

- Their interests align greatly with our work because we focus on training models that are out-of-core (i.e. on multiple machines), and this property allows those models to solve more complicated problems, such as feature identification in images of higher resolutions.
- Their funding does not typically exceed \$500,000 in a year given their award history <https://www.nsf.gov/awardsearch/advancedSearchResult?ProgEleCode=099Y&BooleanElement=Any&BooleanRef=Any&ActiveAwards=true&#results>
- So far, many of their recipients have not recorded more than two years of funding.

## NSF PPoSS

- The main focus of the NSF PPoSS program is also academic institutions, many of the Universities awarded have been graduate students and university faculty, mainly in departments related to computer science and electrical engineering.
- The PPoSS program is funding for a target of systems that are scalable or promote scalability, as to account for the complexity of problems and their needed solutions in the future.
- Their interest can align with our work because our work should be able to scale with the number of machines available as well as their capacities at the time of training. Our splitting of out-of-core models should work with any number of machines as well as any type of GPU that meets our training requirements.
- Their funding is mostly between \$100,000 and \$500,000 total, but they have awarded recipients with \$1,000,000 within the scope of 1 year. A lot of their grants do not tend to be continuous (spanning over multiple years). You can find their recipient history here : <https://www.nsf.gov/awardsearch/advancedSearchResult?ProgEleCode=042Y&BooleanElement=Any&BooleanRef=Any&ActiveAwards=true&#results>

## NSF SPX

- The main focus of the NSF SPX program is also academic institutions, many of the Universities awarded have been graduate students and university faculty, mainly in departments related to computer science and electrical engineering.
- The SPX program is seeking projects that revolve around cyberinfrastructure that exploits parallelism in order to achieve more convenient performance results, such as faster execution times.
- The SPX's ideal recipient matches us quite well because our focus revolves around distributed training using model-parallelism. Since we are building a framework that allows users to train large neural networks on multiple GPUs, and this results is meant to reduce training time significantly, we should be counted as exploiting parallelism to a high degree.
- Their funding is usually between \$100,000 and \$500,000, but \$1,000,000+ awards have been offered. A lot of their grants do not tend to be continuous, but that might be because the program itself is fairly new. You can find their recipient history here : <https://www.nsf.gov/awardsearch/advancedSearchResult?ProgEleCode=042Y&BooleanElement=Any&BooleanRef=Any&ActiveAwards=true&#results>

## SBIR Program

- **What are their foci?**
  - As the name suggests, Small Business Innovation Research(SBIR) focuses on startups or small businesses on transforming scientific discovery or risky ideas into commercial potential and/or societal benefit through the development of products or services.
- **What kinds of projects do they typically fund?**

- Usually innovative technology that requires a lot of R&D design, and a comprehensive tech sectors can be found here, <https://seedfund.nsf.gov/portfolio/>.
- **What might it take to be successful in obtaining a grant?**
  - The company is not primarily owned by venture capital operating companies, hedge funds, or private-equity firms.
  - The idea should be innovative and can make a difference to people worldwide or revolutionize an industry.
  - The SBIR/STTR Programs at NSF have no specific topical or procurement focus, but you need to choose a technology topics from their funding list for the proposal, see more details at <https://seedfund.nsf.gov/apply/>
  - You need to submit a three-page "Project Pitch" that outlines the project objectives.
- **How do they align with your current project? How might you have to adjust your work to meet their foci?**
  - Our project has the potential to be commercialized to be used by many companies that need to scale their application to train a larger deep learning model to improve their accuracy needs.
- **Funding availability:**
  - Phase I proposal for up to \$256,000, which will cover at least six (and up to 12) months of work.
    - Estimated Number of Awards: 250 to 300 (per year, pending the availability of funds)
    - Anticipated Funding Amount: \$70,000,000
  - Phase II awardees receive up to \$1,000,000 over two years.

### 2.2.3 Choose Two Agency Programs to Target

- **NSF CSSI**
  - We choose CSSI over CISE because it provides more funding. Maximum of \$600,000 for CSSI, and \$500,000 for CISE.
- **NSF SBIR**
  - We choose the NSF agency. The reason for that is because NSF provides more options that align with our project goal, and more available funding is provided for the applicant. And, the other agencies, such as the Department of Agriculture and Department of Health and Human Services, are not very suitable for the objective of our project.
- **Outline of our plan** (Note: Since both programs are under the supervision of NSF agency, the Review Criteria for both of them are very similar)
  - In terms of "The Intellectual Merit criterion" [5], we plan to highlight several major reasons and key features of our project that have the potential to advance the knowledge in the field of deep learning and distributed, high performance computing.
  - In terms of "Broader Impacts" [5], we will evaluate and demonstrate the potential impacts of our result that can be important to benefit and contribute to the field of deep learning and distributed, high performance computing.

### 2.2.4 Choose One of the Programs as Your First Target

We choose NSF CSSI because their requirements are best aligned with the focus of our project and provide the largest amount of funding we can apply for. Our project can be integrated into cyberinfrastructure as it builds on the top already available high-performance libraries like TensorFlow, PyTorch, and MVAPICH2. This project has both system-level and application level challenges that are found in NSF CSSI funded projects.

### **2.2.5 Reflect on this exercise Write a short reflection.**

- What did you learn?
  - We learn several things: 1) Funding is hard to get. 2) There are a lot of things you need to consider for writing a good proposal, e.g., program's requirement, their vision and principle, what is the plan of your project, how can you convince them your plan can be successful? How to evaluate the successes of your project? and what is the impact of your project,
- What questions do you still have?
  - Interested to know how many people applied this funding each year? and what are the odds of getting approved?
- What additional training, experience, or support do you believe would help you be successful in this exercise in practice?
  - No, I think I have all I need for this exercise.

### **Reference:**

1. NSF CSSI, <https://www.nsf.gov/pubs/2020/nsf20592/nsf20592.htm>
2. NSF CISE, <https://www.nsf.gov/pubs/2020/nsf20591/nsf20591.htm>
3. List of federal agencies offering SBIR programs, <https://www.sbir.gov/agencies-landing>
4. Tutorial:
  - a. <https://www.sbir.gov/>
  - b. <https://www.sbir.gov/tutorials>
  - c. <https://www.sbir.gov/tutorials/preparing-proposal/>
5. NSF SBIR, <https://www.nsf.gov/pubs/2021/nsf21562/nsf21562.htm>